

Investing in early childhood care and education : the impact of quality on inequality

Citation for published version (APA):

Bauchmüller, R. (2013). *Investing in early childhood care and education : the impact of quality on inequality*. [Doctoral Thesis, Maastricht University]. Boekenplan. <https://doi.org/10.26481/dis.20130523rb>

Document status and date:

Published: 01/01/2013

DOI:

[10.26481/dis.20130523rb](https://doi.org/10.26481/dis.20130523rb)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Investing in
Early Childhood Care and Education:
The Impact of Quality on Inequality

© 2013 Robert Bauchmüller

All rights reserved. No part of this publication may be reproduced, stored in retrieval system, or transmitted in any form, or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing, from the author.

ISBN 978 90 8666 304 0

Cover picture by Paola Hagen-Zanker

Published by Boekenplan, Maastricht NL

Investing in Early Childhood Care and Education: The Impact of Quality on Inequality

DISSERTATION

To obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus Prof. dr. Luc L.G. Soete
in accordance with the decision of the Board of Deans
to be defended in public on Thursday May 23, 2013, at 14:00 hours.

by Robert Bauchmüller

May 2013

Supervisors: Prof. Dr. Chris De Neubourg
Prof. Dr. Daniele Checchi, University of Milan, Italy
Prof. Dr. Gøsta Esping-Andersen, Universitat Pompeu Fabra,
Barcelona, Spain

Assessment Committee:

Prof. Dr. Lex Borghans (Chairman)
Prof. Dr. Nabanita Datta Gupta, University of Aarhus, Denmark
Prof. Dr. Jaap Dronkers
Dr. Paul Jungbluth

Table of Contents

TABLE OF CONTENTS	I
PREFACE.....	V
ACKNOWLEDGEMENTS	VII
LIST OF FIGURES	IX
LIST OF TABLES	X
CHAPTER I - INTRODUCTION.....	11
1. BACKGROUND	11
2. RESEARCH QUESTIONS	13
3. CONTRIBUTION OF THE DISSERTATION.....	16
4. IDENTIFICATION STRATEGIES	16
5. STRUCTURE OF THE DISSERTATION	18
5.1. <i>STUDY I/III</i> : CENTRE-BASED VERSUS HOME-BASED CHILDCARE	18
5.2. <i>STUDY II/III</i> : POTENTIAL GAINS FROM CHILD-CENTRED EARLY CHILDHOOD EDUCATION	19
5.3. <i>STUDY III/III</i> : LONG-RUN BENEFITS FROM UNIVERSAL HIGH-QUALITY PRESCHOOLING	19
CHAPTER II - CENTRE-BASED VERSUS HOME-BASED CHILDCARE (<i>STUDY I/III</i>)	21
1. INTRODUCTION	21
2. LITERATURE DISCUSSION	23
3. DATA DESCRIPTION	27
4. METHODOLOGY	29
4.1. NON-LINEAR LINK BETWEEN CHILDCARE SPELL AND CHILD OUTCOMES.....	30
4.2. INSTRUMENTAL VARIABLE (IV) APPLICATION	34
4.3. STRUCTURAL EQUATIONS MODELLING (SEM)	35
5. ESTIMATION RESULTS.....	37
5.1. OLS ESTIMATES	38
5.2. IV ESTIMATES.....	42
5.3. SEM ESTIMATES	44
6. CONCLUSIONS.....	47
APPENDIX II	49

CHAPTER III - POTENTIAL GAINS FROM CHILD-CENTRED EARLY CHILDHOOD EDUCATION (STUDY II/III)	53
1. INTRODUCTION	53
2. EARLY CHILDHOOD EDUCATION POLICIES IN THE NETHERLANDS	56
3. LITERATURE DISCUSSION	62
4. DATA DESCRIPTION	65
5. METHODOLOGY	69
5.1. ACCOUNTING FOR MISSING TREATMENT INFORMATION	70
5.2. GROUP DIFFERENCES	72
5.3. LINEAR REGRESSION	73
5.4. PROPENSITY SCORE MATCHING (PSM)	75
5.5. INSTRUMENTAL VARIABLE (IV) APPLICATION	77
6. ESTIMATION RESULTS	80
6.1. OLS ESTIMATES	80
6.2. PSM ESTIMATES	83
6.3. IV ESTIMATES	84
7. CONCLUSIONS	87
APPENDIX III	90
DATA ANNOTATION	99

CHAPTER IV - LONG-RUN BENEFITS FROM UNIVERSAL HIGH-QUALITY PRESCHOOLING (STUDY III/III)	101
1. INTRODUCTION	101
2. BACKGROUND AND RELATED LITERATURE	102
3. DATA AND INSTITUTIONAL FRAMEWORK	104
3.1. DATA	104
3.2. ORGANIZATION OF DAYCARE AND PRESCHOOL IN DENMARK	104
3.3. ATTENDANCE RATES AND PARENTAL SELECTION	105
4. PRESCHOOL QUALITY	106
5. METHODOLOGY	110
6. ESTIMATION RESULTS	114
6.1. OLS ESTIMATES	114
6.2. IV ESTIMATES	121
6.3. SIZE OF EFFECTS	127
7. DISCUSSION AND CONCLUSION	128
APPENDIX IV	130

DATA ANNOTATION	137
CHAPTER V - CONCLUSIONS	138
1. MAIN FINDINGS.....	138
2. INTERPRETATION OF RESULTS.....	140
3. POLICY IMPLICATIONS.....	143
4. LIMITATIONS	145
5. CONCLUDING REMARKS.....	147
REFERENCES	149
EXECUTIVE SUMMARY	169
NEDERLANDSE SAMENVATTING	172
DEUTSCHE ZUSAMMENFASSUNG	175
SLOVENŠČINA POVZETEK	178
CURRICULUM VITAE.....	181
MAASTRICHT GRADUATE SCHOOL OF GOVERNANCE DISSERTATION SERIES.....	182

Preface

[to be added]

Acknowledgements

The dissertation is based on three separate studies that are in the process of getting submitted to working papers series or journals in their respective fields. I would like to acknowledge and express my appreciation to the following persons and organisations for their valuable contribution and advice in the preparation of those studies; of course, I am solely responsible for any errors that remain.

In the first place, I would like to thank my supervisors Daniele Checchi and Gøsta Esping-Andersen for their valuable feedback on many versions of my papers and guidance and support in organising my research. And I would like to thank the team of the Graduate School of Governance at Maastricht University, in particular Chris de Neubourg and Mindel van der Laar, for the support and fellowship grant that allowed me to carry out this research.

I would like to thank the European Centre for Analysis in the Social Sciences (ECASS) for a grant funding my research visit to the Institute for Social & Economic Research at the University of Essex, UK, the EqualSoc network for a grant funding my research visit to the Department of Economics, Business & Statistics at the University of Milan, Italy, and the Centre for Strategic Educational Research (CSER) for a grant funding my research visit to the Danish Institute of Governmental Research in Copenhagen, Denmark.

With respect to my first study (chapter II), I would like to thank the Data Archiving and Networking Services' Electronic Archiving System (DANS-EASY) for granting me access to the PRIMA dataset *P1701*. In this regard, I also acknowledge the excellent data collection implemented by ITS in Nijmegen and the SCO-Kohnstamm Institute in Amsterdam and the data documentation written by Driessen et al. (2006).

Access to the data used for my second study (chapter III), as well as the opportunity to contribute to its field-work design, has kindly been provided by Paul Jungbluth and the KAANS Research Centre at Maastricht University. I thank the municipalities, daycare centres, preschool kindergartens, elementary schools and parents in South Limburg who have participated and cooperated in the data collection, as well as Lex Borghans who helped me identify this research opportunity, and match postcodes to GPS coordinates. The paper evolved from an earlier version that improved through comments received at the EDULEARN10 conference in Barcelona. I would also like to thank Jessica Hagen-Zanker and Elma Nap-Kolhoff for their valuable comments.

The third study (chapter IV) is based on joint research with Mette Gørtz (Department of Economics, University of Copenhagen) and Astrid Würtz

Rasmussen (Department of Economics and Business, Aarhus University). I have learned a lot from our good cooperation and am very grateful for the research that we have done together. I would like to thank the Centre for Strategic Research in Education and Competence Building for their financial support of our research. The Danish Institute for Local and Regional Government Research (KORA; former Danish Institute of Governmental Research, AKF) kindly hosted my research visit to Copenhagen and assisted us accessing online data sources of Statistics Denmark; later on, the Danish National Centre for Social Research and the Department of Economics at the University of Copenhagen generously granted me an account and support to access these data sources from abroad. We appreciate comments on previous drafts of the paper by Virginia Maestri, Marianne Simonsen, participants of the AKF Quantitative Seminar, the NCoE Workshop Bornholm, the Annual ESPE Meeting in Essen, the EPC Conference in Vienna, the annual CSER meeting in Copenhagen (all in 2010), as well as the Joint ASB/IAB Workshop, the SFI Advisory Research Board Conference and a seminar in AKF (all in 2011), and three anonymous reviewers.

Finally, I would like to thank the assessment committee (Lex Borghans, Nabanita Datta Gupta, Jaap Dronkers, and Paul Jungbluth) for having swiftly read my dissertation and providing helpful suggestions to further improve it.

List of Figures

Figure I-1: Schematic early childhood care and education options studied	15
Figure II-1: Link between child outcomes and centre-based childcare experience .	32
Figure II-2: Distribution of standardized child outcomes	49
Figure III-1: Distribution of standardized child outcomes	90
Figure III-2: Calculated weights.....	91
Figure III-3: Distances to closest preschools without child-centred ECE.....	97
Figure III-4: Distances to closest preschools with child-centred ECE.....	98
Figure III-5: Closest preschool with child-centred ECE over alternatives	98

List of Tables

Table II-1: Summary statistics on child outcomes by childcare arrangement	28
Table II-2: OLS regression output.....	39
Table II-3: 2SLS regression output, applying the instrumental variable	42
Table II-4: SEM 1 st Stage - <i>probit</i> centre- versus home-based childcare choice	44
Table II-5: SEM 2 nd Stage - OLS regression model including Mills ratio.....	46
Table II-6: Descriptive statistics	51
Table II-7: Variation of average childcare attendance at each province	52
Table II-8: Simultaneous quantile regression.....	52
Table III-1: Summary statistics on child outcomes.....	72
Table III-2: OLS regression output	80
Table III-3: Propensity score matching	83
Table III-4: 2SLS regression output, applying all 4 instrumental variables	85
Table III-5: Test scores in five categories and averages of standardized scores.....	91
Table III-6: Descriptive statistics	92
Table III-7: Length of different childcare experiences	93
Table III-8: Appropriate parental quoting of <i>Speelplezier</i> ECE method	94
Table III-9: Univariate sample comparison, with and without ECE information....	94
Table III-10: Multivariate probit regression for different samples.....	95
Table III-11: Preschool comparison, with and without child-centred ECE.....	96
Table III-12: Distance information between homes and preschools	97
Table IV-1: Child and preschool population in 1998	106
Table IV-2: Quality characteristics.....	109
Table IV-3: OLS regression output including all control variables.....	116
Table IV-4: Simultaneous quantile regression	120
Table IV-5: 2SLS regression output, applying all 8 instrumental variables.....	123
Table IV-6: Descriptive statistics of control variables	130
Table IV-7: Danish grades and the averages of their standardized scores	130
Table IV-8: Aggregated preschool quality averages across Danish counties	131
Table IV-9: Quality difference of conservative municipalities, 1993 election	132
Table IV-10: Municipal child population aged 3, %-change 1995/96 by centile	132
Table IV-11: Availability of guaranteed preschool place scheme in municipality	133
Table IV-12: OLS regression output including all control variables (<i>full output</i>)...	134

Chapter I - Introduction

1. Background

Public investments in early childhood care and education provisions (henceforth also referred to as childcare or daycare) have increased significantly in recent decades, in both developed and developing countries. Such quantitative extensions take, for example, the form of lowering compulsory schooling ages, introducing rights to a place in daycare for children at earlier ages and broadening the coverage of provisions to wider groups of children. The extension of non-parental childcare supply follows a massive growth in demand for those services.

Three major trends are promoting demand for childcare services. Firstly, female emancipation in the labour market has increased maternal employment, resulting in an increased need for non-parental childcare (cf. Bianchi, 2000). Secondly, while having seen considerable GDP growth since World War II, many countries still report high or even increasing child poverty rates (UNICEF, 2005), whereas inequalities in child outcomes emerge already in the early years (cf. Bradbury, Corak, Waldfogel, & Washbrook, 2011). Children with single parents and without income earners in their household are especially exposed to the risk of growing up and being trapped in disadvantageous environments, requiring public investments to guarantee equality of opportunity (UNICEF, 2000). And thirdly, economic progress has been accompanied by significant demographic changes, reducing the number of children born. The resulting imbalance in the age structure, paired with the need for a well-trained workforce for the future's knowledge-intensive economies, urges governments to make investments that assure that no child falls behind.

In addition, early childhood is increasingly recognised as the crucial period to stimulate child development and thus improve chances to do well in later schooling and the labour market (cf. Knudsen, Heckman, Cameron, & Shonkoff, 2006). Public investments in early childhood provisions are expected to lead to substantial social returns (cf. Cleveland & Krashinsky, 2003; Dickens, Sawhill, & Tebbs, 2006; Esping-Andersen, 2005; Rolnick & Grunewald, 2003).

A few countries, such as, for example, Denmark, have already approached a near universal childcare coverage for children from a few months after birth; others, like the Netherlands, are following swiftly. Yet, many countries still do not manage to provide the quantity and quality of childcare provisions that are deemed necessary to address the increased demand. Stronger orientation in most countries to fiscal consolidation and performance-oriented budgeting requires that funds are addressed to the most effective social investments. A lack of sufficient evidence on the importance and adequate quality of public investments in childcare provisions

reduces the readiness of policymakers to allocate sufficient public funds to the childcare sector. At the same time, policymakers try to serve the excess demand for childcare services by increasing the quantity of childcare places rather than promoting the quality of services. However, evidence on the effectiveness of childcare provisions is particularly weak with respect to quality aspects. This makes it more likely that public childcare investments turn out to be ineffective, delivering suboptimal social returns.

Nevertheless, the strong theoretical case that is made for early childhood investments generates growing attention for quality aspects of childcare investments. Early investments are expected not only to improve the development of innate productivity, but also to have a multiplier effect on later human capital investments (Heckman, Krueger, & Friedman, 2002)¹. Contributions from different research fields have led to a growing consensus that skills are a result of both nature *and* nurture, genetic predispositions being triggered by the interaction with the environment. And neurobiological research has shown that the plasticity of the brain is greatest during early childhood. These discoveries re-emphasized the importance of having an adequately stimulating environment in that period. Cunha et al. (2006) argue that innate abilities of a child might be fully developed already before entering compulsory schooling and thereafter would be only amendable at high costs and efforts.

For children from disadvantaged backgrounds, childcare provisions are supposed to lower the risk of falling behind throughout their schooling, and thus breaking the links of inherited inequality (cf. Björklund & Jäntti, 2008; Bowles & Gintis, 2002; Bowles, Gintis, & Osborne Groves, 2005). Policies assuring quality childcare intend to promote homogenizing early childhood investments (Esping-Andersen, 2006). And efficiency and equity of education are increasingly considered as mutually reinforcing (cf. EC COM, 2008; Leseman, 2009).

Disadvantages may result, for example, from having only one caring parent in the household, limited economic resources, parents with low educational attainments or belonging to an ethnic minority. Being a boy has also been associated with a higher risk of falling behind. Yet, the literature is not conclusive on what a stimulating environment for such disadvantaged children should look like. Many have argued that policymakers should strengthen the economic conditions of families at the first place. There is significant evidence supporting the importance of sufficient financial resources within the family to invest in the child; i.e. poor families might face a limited access to adequate early childhood provisions due to

¹ This concept is often referred to as 'skills beget skills', 'success breeds success' or 'learning begets learning'.

credit constraints. However, there is also evidence indicating that many families who are financially constrained are likely to underinvest in the child even if more resources were available, which is calling for alternative care provisions (see e.g. Mayer, 1997).

A selected number of longitudinal studies is usually put forward as strong evidence to stress the potential positive impact of non-parental childcare experiences on child outcomes, in particular for children from disadvantaged backgrounds. Evidence, for example, from the Perry Preschool, Syracuse FDRP and Chicago CPC programmes confirms the expectations of positive impacts of childcare investments on employment, income, and reduction of crime and delinquency rates (cf. Campbell, Helms, Sparling, & Ramey, 1998; Schweinhart, Barnes, & Weikart, 1993). However, much of that evidence is confined to small-scale, high-quality, intensive programmes that are targeted at disadvantaged children and that are more typical for Anglo-Saxon countries. Little is known so far about whether it holds also for more universal childcare provisions and other country settings. A meta-analysis by Nores & Barnett (2009) of international (non-U.S.) studies indicates, however, that positive expectations about childcare investments may also be found in other countries.

Acknowledging the theoretical support for public investments in early childhood and seeking to extend the still limited empirical evidence, the dissertation addresses the following questions and hypotheses to inform policymakers on the effectiveness of essential quality aspects and to search for the best responses to the needs of disadvantaged children.

2. Research questions

When starting to look into public investments in childcare provisions, the first question that arises is: (1) *Do centre-based childcare provisions stimulate children better than home-based provisions?* I presume that centre-based childcare should be beneficial because of, for example, exposing the child to professional care staff and environments whose quality is frequently assessed, promoting peer learning and regular stimulating activities, and sharing access to best child-raising practices with parents. Yet, there are also good reasons to doubt benefits of centre-based childcare, too, i.e. that childcare centres may not be able to address individual development needs of each child, that the overall quality of childcare centres may not match the average quality of home-based childcare because of a lack of resources or problems in monitoring quality, and that detachment of children from their parents may be harmful at too early ages. Even if the childcare centre's quality would be sufficient, positive effects may only become visible with more hours being spent in a childcare centre. I expect decreasing returns to more extensive spells, i.e. with an earlier age at which a child attends a centre or with

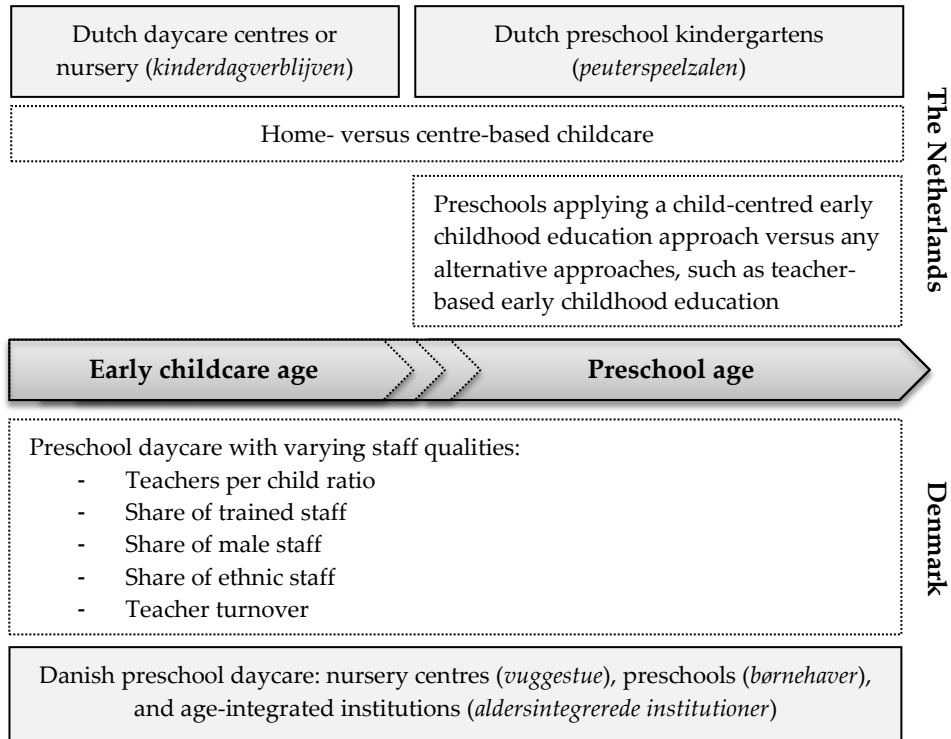
more hours per day the child spends there. The more attention is put at a centre to early education, i.e. at preschool kindergartens, the stronger should be the effects.

Such early childhood education approaches are introduced by a growing number of care providers. Extending schooling principles to preschool ages acknowledges the importance of stimulating learning experiences already in early childhood when child development is very sensitive to such stimulus. Using that logic, some countries start earlier with compulsory schooling than others. However, when deciding on a specific education approach the following question is often put forward: (2) *Are learning experiences, in particular for disadvantaged children, more stimulating if the applied early childhood education approach is more child-centred rather than teacher-based?* I presume that child-centred education approaches should be more beneficial as they are more holistic in their understanding of and response to individual needs. However, a child-centred education approach might require sufficient public investments to allow childcare staff to devote enough time to each child. Accordingly, there is some doubt that childcare arrangements, which are provided at a near universal scale, can assure suitable development stimulation of individual children. Also, child-centred approaches may crucially depend on childcare staff having the appropriate skills to understand individual children's development needs and how to address them, i.e. because of having received insufficient training.

The biggest budget line of any childcare investment is typically spent on the childcare staff (often referred to as teachers). The way funds are invested in childcare staff may be strongly correlated with the impact of the resulting childcare experiences. This raises the following question: (3) *Does childcare with more staff members per child, more professional teachers, lower turnover of staff, more male staff and staff from ethnic minorities lead to better child outcomes?* More staff, less frequent staff changes and better trained staff supposedly allow more quality time with individual children and a strong teacher-child interaction. Even though, some doubt that quite substantial changes in the staff would be needed to make a difference, whereas such changes would be too costly as compared to other potentially beneficial options. Male and ethnic minority staffs are usually underrepresented at childcare institutions, whereas boys and ethnic minority children tend to have higher risks of falling behind in their schooling. Increasing the shares of male and ethnic minority staff is expected to draw more attention to particular needs of boys and ethnic minority children. However, there are also reasons to believe that more ethnic staff may hinder integration of minority children, i.e. because of making it more likely that ethnic minority children can use their native languages. And there are some doubts about the importance of male care staff that are linked to the importance of gender-specific factors influencing child development.

The following figure sketches aspects of early childhood care and education studied in this dissertation.

Figure I-1: Schematic early childhood care and education options studied



Besides being interested in the overall effects of such quality aspects of childcare investments, policymakers hope to make the best investments to promote equality of opportunities. (4) *Do disadvantaged children benefit above average from specific childcare quality characteristics?* The logical expectation is that centre-based childcare should deliver positive impacts at least for the most disadvantaged children, as they are likely to receive less stimulating care at home. Similarly, child-centred early childhood education should be able to better address the needs of individual disadvantaged children – teacher-based education approaches might be too challenging if children start with developmental delays. With respect to the staff quality aspects, disadvantaged children are expected to benefit from a higher number of staff, with better training and less turnover, as those factors contribute to a more personal, professional and constant care of individual children. A more diverse staff in terms of male members and ethnic minority backgrounds is likely to contribute to better development of boys and children from ethnic minorities respectively. Yet, critics of stimulating child development through universal childcare provisions often stress that even high-quality childcare arrangements may be insufficient to compensate development risks if disadvantaged children

spends only a limited share of their time in centre-based childcare, and that - overall - possible benefits for the disadvantaged may be outweighed by insufficiently stimulating advantaged children.

3. Contribution of the dissertation

The expectations about positive effects of childcare investments are considerable. Yet, the supporting empirical evidence is still thin, in particular on the importance of different quality aspects that policymakers may influence through their investment decisions. This dissertation adds new evidence on three areas of childcare quality for two European countries that devote extensive public investments to childcare provisions - the Netherlands, an example of a country that has recently accelerated substantially the extension of its childcare provisions, and Denmark, an example of a country that has already reached near-universal high-quality childcare provisions.

The case of the Netherlands is studied with data from a representative national cohort study of the year 2004/06, adding to earlier analyses by using a new research design. Additional light is put on the Dutch case by using an analogous follow-up study on all children in the southern part of the province of Limburg, collected at more than 200 elementary schools in 2009/10. The Danish case is studied jointly with two colleagues from Danish research institutions, using a unique administrative dataset that integrates childcare, personnel and citizen registers of the years 1995 to 2008. The preparation of the last two datasets for this dissertation contributes new opportunities for researchers to study early childhood investments on a number of short- and long-term child outcomes.

4. Identification strategies

Besides the availability of good data, the study of the effects of childcare quality characteristics largely depends on a good research design to identify true causal effects. The best possible design would allow a random allocation of childcare quality aspects among a sufficiently large group of children, so that all mechanisms of treatment allocation would be controlled by the researcher. However, such experiments are typically not applicable for ethical reasons, e.g. problematic exclusion of some people from a presumably good treatment, or practical reasons, e.g. possible failure to contain external influences. Beyond that, controlled experiments are often subject to some form of selection bias and are very sensitive to drop-outs when they have a longitudinal character. Yet, the evaluation literature builds on some well-known controlled experiments such as, for example, the before-mentioned Perry Preschool, Syracuse FDRP and Chicago CPC programmes. The experimental character of those programmes allows the drawing of conclusions about the impact of childcare investments from the difference in

average child outcomes between the treatment and control groups (see, for example, Heckman, Moon, Pinto, Savelyev, & Yaditz, 2010).

However, experiments are usually applied for highly targeted, intensive, specialised programmes; despite strong internal validity they lack external validity. To measure effects on a larger scale requires data collected at a broader level, e.g. administrative registries or surveys. In most cases, such data takes the form of a cross-sectional collection of information at a certain point of time; however, some studies have also longitudinal characters that repeatedly observe the child and his background at various ages. Cross-sectional and longitudinal data do not allow a controlled treatment allocation *ex-post*; researchers typically study them by some type of correlational study design – child outcomes are compared across groups under control of rich sets of background characteristics. Causal inferences of such studies rely on the assumption that all relevant determinants of the child outcome are observed. This tends to be a weak assumption as in many cases we cannot be sure that we cover the full correlational model and its correct functional form, account for all relevant factors, or are able to exclude the presence of endogeneity problems.

Results may be biased because some unobserved variable might have influenced the selection of the sample that we studied – the studied child outcomes do not represent the whole population. Or we may face a bias because an independent variable included in the model is potentially a choice variable, correlated with unobservable variables relegated to the error term. A better control for unobserved factors may be possible when the longitudinal character of a study allows some form of fixed effects or differences-in-differences design, using earlier (pre-treatment) observations of the child to compensate possible biases.

Other methods, for example, non-parametric approaches such as, for example, propensity score matching allow us to relax some assumptions about the functional form of the estimation model, which describes the correlations between outcomes, treatments and the control set. In light of the many assumptions that need to be made about the one or the other model form and estimation strategy, it has become popular to model the allocation mechanisms theoretically. A choice model is used as control function for estimating causal treatment effects. Yet, the problem remains that factors and mechanisms influencing the treatment allocation are unobserved or not accounted for, hence our estimates may be biased. Sensitivity studies may allow an assessment of the direction and size of such a bias, but might not minimize it completely. In that case, quasi-experimental research designs are recommended, implying the identification of any external source of variation that influenced the probability to choose a specific childcare provision but is unrelated to the assessed child outcomes.

Potential candidates of quasi-random treatment allocation are policy reforms that cause, for example, changes in the welfare rules that determine the access to specific childcare provisions but do not influence the child directly. Such quasi-experimental information on the treatment allocation may be used as instrument for the treatment variable, as exclusion restriction in a choice model or as an aid within a discontinuity design. A general problem of this approach is that there are rather few convincing external sources of treatment allocation that are fully detached from parental choices and not influencing the child in another way. Even if a variable is supposed to be a valid source of randomness, it might provide ‘too’ weak information on the treatment variation or be useful only within a certain local range of the treatment variable.

In this dissertation, I apply a set of such designs, provisional on what is most applicable and adequate for the available dataset and given research question to be studied (for further information on such designs refer to, for example, Baumüller, 2008; Haveman & Wolfe, 1995; Manski, 1993; Todd, 2008).

5. Structure of the dissertation

This dissertation consists of three main chapters, II-IV, which represent self-contained studies, addressing the research questions mentioned above in Section 2. The chapters explore the effects of quality aspects of early childhood care and education aspects on short- and long-term outcomes in the Netherlands and Denmark, using three unique datasets, two of which have been made available to researchers for the first time. In order to understand the heterogeneity of the estimated effects across children of different backgrounds, the analyses include various interaction terms to account for subgroup effects, for example, for children from poor households, single parents, low-educated parents and boys versus girls.

5.1. *Study I/III: Centre-based versus home-based childcare*

Centre-based childcare is seen as public investments to facilitate maternal employment. Recent theoretical research proposes that such investments potentially lead to substantial gains in child development and thus to high returns for the society as a whole. However, the empirical evidence is still scarce and often contradictory. I study rich survey data of a large-scale cohort study of children living in the Netherlands at the beginning of the new millennium. The Netherlands has made substantial investments in the last two decades to make the market of centre-based provisions more professional and far-reaching and to improve children’s school readiness. I study the impact of experiencing centre- rather than home-based childcare on language, cognitive and non-cognitive development, assessed at the age of 6. To assess whether very long or intensive childcare spells can be harmful, I account for possible non-linearity in the correlation between the

centre-based childcare experience and the child outcomes. As sensitivity analyses, I also apply instrumental variable and structural equation modelling approaches to try to correct for potential biases in my estimates that would result, for example, from unobserved heterogeneity of parents and children. For both ordinary least square estimates as well as the sensitivity analyses the results do not support the significant short-term effects of centre-based childcare stated in the literature.

5.2. *Study II/III*: Potential gains from child-centred early childhood education

Early childhood education programmes are presumed to have positive effects, in particular for children who are at risk to fall behind during their schooling. However, there is disagreement on whether such programmes should be more teacher and curriculum based or instead centred on the individual child. In this chapter, I study child-centred education programmes that are used at preschools in the Dutch province of Limburg, which is in fact mainly a study of ‘Speelplezier’, a new child-centred programme which has recently been certified as being ‘in theory’ effective in raising children’s school readiness but has not been evaluated yet. I use a rich dataset covering the first three grades at elementary schools in South Limburg for the year 2008/09 to evaluate the impact of child-centred versus alternative early childhood education options. I estimate ordinary least squares effects of attending a preschool, applying child-centred education approaches onto test scores from the beginning of elementary schooling. I control for alternative childcare experiences and various child- and family-related characteristics and re-weight observations of the studied sample to represent population averages. I argue that access to a preschool kindergarten applying child-centred education is to some degree exogenously determined. In a further effort to confirm the causality of effects, I also use propensity score matching and instrumental variable estimation techniques. The different estimate techniques show that child-centred early childhood education has not yet led to significant effects on child development.

5.3. *Study III/III*: Long-run benefits from universal high-quality preschooling

This study investigates the role of preschool quality in children’s performance at the end of primary school (age 16). We use a unique dataset based on Danish administrative registers where children are linked to their preschool and its teachers. Based on this, we generate five main quality indicators of preschools. Controlling for child background factors, we study the effects of those preschool quality indicators on long-term outcomes. The main child outcomes that we study are written Danish test scores from final (9th grade) elementary school exams, referring also to similar analyses with other language and cognitive outcomes. To

address the potential homogenizing effect of universal high-quality preschooling, we also study the heterogeneity of effects across different subgroups of children who have higher risks of falling behind in their schooling trajectory. A sensitivity analysis using an instrumental variable approach is used to address possible biases that may result from unobserved variables in the ordinary least square estimates. The different analyses show that four out of the five quality indicators are associated with significant improvements in children's test results in Danish - albeit the relationship being numerically modest. Boys benefit more from preschool quality than girls and ethnic minority children benefit from higher staff stability. Instrumental variable estimates do not produce more consistent estimates.

Chapter V summarizes and discusses the results of the three studies and stresses policy implications. The chapter concludes the dissertation with a discussion of its limitations and possible next steps in advancing the research, as well as some concluding remarks.

Chapter II - Centre-based versus home-based childcare (study I/III)

1. Introduction

Centre-based childcare provisions are extended in most countries. The primary objective of this extension is to increase female labour market participation. However, in doing so policymakers are often beset by worries about potential harm for their children if they attend such centre-based childcare arrangements, at a very early age or at institutions of insufficient quality. Concerns include a lack of sufficient evidence whether centre-based childcare provisions provide an environment that sufficiently stimulates child development by replacing the time spent with the child at home and whether they can provide an extraordinary chance for disadvantaged children to develop better than by being cared for 'only' at home.

The Netherlands is an example of a country with a long history of improving and extending childcare provisions. The number of centre-based childcare places increased between 1989 and 2004 from 20,000 to 200,000. The main providers are daycare centres and preschool kindergartens.² Early childhood care is provided by daycare centres (or nurseries) that serve children from the age of about 6-8 weeks up to the end of elementary schooling, which includes out-of-school care provisions.³ Preschool kindergartens cater to the age group of 2-4 years and follow the same basic childcare quality standards. However, there is substantial variation in quality levels across those two types of provisions as well as within them.

Preschool kindergartens are regarded as having a higher average quality as they have focused earlier on professionalization of staff and strongly on stimulating child development; many of these use an early childhood education approach. The child population at preschools tends to come from more disadvantaged backgrounds. Subsidies targeted at those disadvantaged children are used to provide them with an additional education stimulus in preschools. In most cases municipalities subsidize, for example, an additional third and fourth half-day of weekly preschool attendance to children attending preschools that run a certified early childhood education programme. Such programmes have been introduced to a growing number of preschools – first only to urban areas, but increasingly also to rural areas. The mean quality of preschool kindergartens thus tends to be more

² The Dutch names are '*kinderdagverblijven*' and respectively '*peuterspeelzalen*'.

³ Dutch elementary schools are called '*basisscholen*' and cover the grades 1-8 (ages 4-12).

professionalised, e.g. in terms of the training level of staff. Thus a preschool experience is likely to have a bigger impact on child outcomes than an ordinary daycare experience. Yet, recent reforms and investments in the daycare sector address the quality differences between both types of childcare providers.⁴

Early childhood care and education investments are placed by Cunha and Heckman (2007, 2008) in a dynamic model, in which skills form in multiple stages throughout childhood. Acknowledging also the multiple dimensions of child development and the malleability of abilities at early ages, they argue that earlier investments produce multiplicative skill effects throughout later stages. They argue that non-cognitive skills are most elastic to parental inputs and stronger initial non-cognitive skills promote cognitive development at later stages. Yet, it is important to know which childcare investments have the potential to produce multiplier effects, which implies looking first at the direct outcomes at school enrolment to understand the initial gains.

However, there is still little evidence available on the effects of either source of childcare – be it home- or centre-based – for children from well-off backgrounds as well as for those disadvantaged children who are expected to benefit above average from good quality centre-based care arrangements. Despite a growth of the international evaluation literature on non-parental childcare provisions, a majority of studies focuses on small-scale and targeted childcare programmes rather than on large-scale provisions, such as those that can be found, for example, in the Netherlands. Evidence is often limited in its validity to specific groups of children or country settings.

This chapter assesses whether attending centre-based childcare is related to any significant development gains by the time children start elementary school as compared to home-based care provisions, and whether disadvantaged children benefit above average. It considers daycare and preschool kindergarten jointly as centre-based childcare (treatment group) while home-based care refers to any alternative care provisions (control group), including primarily parental care but also home-based care for small groups of children by a child-minder. In 2004/05, newly enrolled children at elementary schools had attended on average about 98 half-day sessions at preschool kindergartens and 84 half-day sessions at daycare centres, which aggregates to about 182 half-day sessions in any centre-based childcare arrangement.

⁴ The law 'Wet Kinderopvang' of January 1, 2005, included, next to the reorganization of financing procedures, a new regulation for quality standards, and has the intention to extend and improve the childcare sector so as to allow more mothers to work.

I use data from the national cohort study PRIMA and follow up on an earlier, related study by Driessen (2004). New additions as compared to Driessen's paper are the application of a different estimation strategy as well as the use of a more recent wave of the PRIMA data collection. Using a pseudo-experiment and a care selection model, I attempt to identify causal effects of attending daycare centres and preschool kindergartens, as compared to any alternative care arrangements that are home-based. Child outcomes are assessed at the age of six (middle of 2nd grade) and measure cognitive, non-cognitive and language development. To address the question of how much childcare is beneficial and when it turns out to be harmful, I account for non-linearity in the relationship between the duration of care attendance and child outcomes – accounting for attendance information on both types of centre-based childcare separately.

A review of international and Dutch childcare evaluation literature (Section 2) is followed by a description of the analysed dataset (Section 3) and the applied empirical strategy (Section 4). Thereafter, the estimation results are presented and discussed (Section 5), and the chapter is rounded off with final conclusions and policy recommendations (Section 6).

2. Literature discussion

A number of studies provide a comprehensive overview of evaluations of childcare investments that shall compensate in particular for parental childcare time or disadvantaged family backgrounds (see, for example, Camilli, Vargas, Ryan, & Barnett, 2010; Cunha et al., 2006; Heckman, 2008; Karoly, Kilburn, & Cannon, 2005; Nores & Barnett, 2009; Waldfogel, Han, & Brooks-Gunn, 2002). The evidence in those studies suggests positive effects in particular when children with disadvantaged backgrounds benefit from more extensive, targeted, high-quality interventions (see, for example, Esping-Andersen, 2004).

Estimated effects of childcare investments are particularly strong for cognitive development domains; even so, other domains such as non-cognitive skills are indicated to be equally relevant for long-term success in life (see, e.g. Borghans, Duckworth, Heckman, & Weel, 2008; Fletcher, 2012; Heckman, 2008; Heckman, 2011; Heckman & Masterov, 2007). Effect measures are lower when large-scaled childcare provisions are evaluated, and often positive effects are shown to be non-lasting.

Magnuson, Ruhm and Waldfogel (2007) estimate the impact of prekindergarten attendance on children's school readiness in the U.S. They find that prekindergarten attendance is positively associated with cognitive and language but negatively with non-cognitive outcomes. Effects on non-cognitive outcomes are more persistent at the beginning of the school trajectory, whereas the cognitive gains are more persistent for disadvantaged children. In a later study on U.S. data,

Magnuson and Waldfogel (2005) study the heterogeneity of effects of early childhood care and education attendance on school readiness across different ethnic, racial and economic backgrounds. They find that childcare attendance is associated with better school readiness and recommend that a substantial increase in enrolment of Hispanic, Black and poor children, as well as in the quality of the childcare that those children attend could potentially improve school readiness.

Gregg et al. (2005) study effects of the large increase in maternal employment in the UK in the last 20 years. Besides studying adverse family background factors that could have negative implications on child cognitive outcomes, they look at the effects of the use of non-maternal childcare on cognitive school readiness. They find that such care can replace the lack of maternal care if the quality is sufficient and recommend that affordable childcare should be available, particularly for very young children. Brilli, Del Boca & Pronzato (2011) confirm such findings by studying the case of Italy. They find positive effects of public childcare, in particular on children's language development and for children with low educated mothers in poor regions. In contrast, Bernal (2008) finds in a comparable analysis with U.S. data that very early maternal employment and use of childcare can have sizable negative effects, in particular for children with higher ability endowments. This reflects that the literature is still inconsistent in the effect estimation of childcare attendance.

The large majority of evaluations of home-based versus out-of-home childcare solutions do not differentiate by the type or quality of non-parental childcare solution. An important contribution has been a study by Datta Gupta and Simonsen (2010) in which they used rich Danish administrative and survey data to assess the effects of home-based care versus preschool and family daycare on a non-cognitive child outcome at age seven. While they find no significant overall differences, they find negative effects of family daycare attendance for boys whose mothers achieved only lower education. In a follow-up study, they assess a non-cognitive child outcome at the age 11 and find no significant differences between types of daycare (Datta Gupta & Simonsen, 2011). Using the same data, Esping-Andersen et al. (2011) compare high-quality centre-based care versus lower-quality child-minding versus family care in Denmark and the U.S. They find a significant and positive effect of high-quality care on reading skills at the age of 11 in Denmark, especially for disadvantaged children. For the U.S. they show that formal school- or centre-based care relates to significant cognitive development gains at school entry. However, they show that these effects are not lasting, in particular for disadvantaged children, e.g. because they may find themselves later on in low-quality schools that undo the initial stimulus effects.

Notwithstanding the growing international research on the effects of childcare, Dutch research on such investments is still limited. Studies have usually been

small-scale. The development of pedagogically more structured programmes at preschools has predominantly been accompanied by studies on how much the programmes fulfil official quality standards rather than on how they actually affect the individual child. There is no evidence of daycare centres having significant effects, but there are some indications that preschool attendance may have some positive effects on language and cognitive skills (Van der Vegt, Studulski, & Klopprogge, 2007).

There are two extensive national empirical studies on the effects of attending early childhood care and education. Driessen (2004; see also Driessen & Doesborgh, 2003) use several earlier waves of the Dutch large-scale PRIMA cohort study (1996-2000) to analyse the (co)variances between different early childhood interventions and child language, cognitive as well as non-cognitive outcomes. The analysis controls for various child and family characteristics and cannot confirm any effects of Dutch early childhood care and education investments on the tested child outcomes, either in the short term (test scores at age 6) or in the medium term (test scores at ages 8, 10 or 12). As soon as child and family background characteristics are controlled for, the weak relationships between early childhood care and/or education participation and test outcomes turn out to be insignificant.

Driessen suggests that insufficient quality of interventions, low intensities and short durations of participation as well as early fading out due to discontinuities of stimulus may be reasons why no significant effects are found. He also quotes Blok & Leseman (1996) in noting that early interventions may not be sufficient to overcome deficiencies in the home environment. However, he points out that additional public investments in early childhood care and education, which have taken place since the year 2000, may eventually lead to significant effects.

The second major evaluation study of Dutch childcare investments was done by Nap-Kolhoff et al. (2008). They use later waves of the PRIMA study (2002-2004) and supplement it with qualitative information on childcare from a web survey of schools and preschools on their use of early childhood education programmes and a number of supportive case studies. Even though Nap-Kolhoff et al. have much more detailed quality information about childcare characteristics available than Driessen and look at early childhood interventions that have progressed in quality since Driessen's study, they also find no significant mean effects, once background characteristics of the child and school fixed-effects are accounted for.

Nap-Kolhoff et al. also look at interaction effects. They find that attendance at a preschool that uses such an early childhood education programme leads to less negative outcomes for several groups: children who are of foreign background (but neither Turkish nor Moroccan), children of native origin, children who have low educated parents, and children who have parents with a middle level of education attainments. While this indicates that the desired stimulus on the target group may

have been reached to some degree, the picture is not fully coherent. There is no difference for children of higher educated parents or for Turkish and Moroccan children of low educated parents. Nap-Kolhoff et al. contemplate that conclusions on mid- and long-term childcare effects cannot be made yet. They also indicate that childcare quality has been lower at locations where such programmes have been implemented more recently. They suggest that more detailed quality information needs to be collected and linked to databases to be able to produce evidence that can properly inform childcare policymaking.

A number of international studies evaluate the importance of childcare duration (years) and intensity (hours per week) in generating an impact on child outcomes. The studies indicate that longer, more intensive childcare experiences are more likely to cause significant development gains, up to a certain amount.

Gorey (2001) finds in a meta-analysis evidence a strengthening of effects with more intense and longer programmes. Another meta-analysis of 117 childcare evaluations by Leak et al. (2010) assesses the importance of the starting age and duration of childcare spells. Their results indicate that a starting age below 3 is associated with higher development outcomes, whereas the duration is not significantly related to child development. Barnett & Lamy (2006) show that childcare as of the age of 3 instead of 4 might increase development gains, e.g. in vocabulary development. Leuven et al. (2010) use the same Dutch PRIMA data and school readiness indicators as used in this chapter; they look at the age variation of elementary school enrolment which usually takes place around age 4 and resembles preschooling in the Netherlands, and show that earlier schooling leads to development gains. A cross-country study using PISA data of Braga, Checchi & Meschi (2011) support these findings. They show that reforms expanding the access to pre-primary education are associated with an increase in average educational attainments.

The intensity and duration of a child's childcare attendance and the socio-economic and ethno-cultural background are determinants of childcare's effectiveness. Landvoigt, Mühler & Pfeiffer (2007) look at two aspects of the length of kindergarten attendance in Germany, duration in terms of years of enrolment and intensity in terms of daily hours. They find that both - non-attendance and full-day attendance - are associated with a significantly lower probability to reach the highest secondary school track. However, they also show that those two groups of children have weaker family backgrounds than children who attend half-days. Also, they indicate that intensity may matter more than duration. Barnett and Lamy (2006) show that longer preschool attendance is associated with higher cognitive and language outcomes. Schütz, Ursprung and Wößmann (2008) show in a comparative study of 54 countries that longer pre-primary education in terms of duration and enrolment is positively associated with cognitive

development at mid-level schooling age; systems with longer pre-primary education spells show more equality of opportunity.

Yet, some studies argue that too intensive and very early attendance at childcare institutions could be harmful for children, in particular if these are of insufficient quality. Evidence shows, for example, that within the first year of life centre-based care can be harmful if its quality cannot compensate for the temporary detachment of new-born children from their parents (see, for example, Belsky & Rovine, 1988). However, they indicate that the potential harm relates to attendance at a very early age rather than very intensive attendance; the harmful effect should diminish with increasing ages. Loeb et al. (2005) find that intensive centre-based care at very early ages has positive effects on cognitive and language outcomes, whereas harmful effects on socio-behavioural outcomes are amplified. Results vary by income level and race - poor and Hispanic children benefit above average in their cognitive and language development. Waldfogel (2004) suggests that early intensive interventions might have a positive effect in particular on the socio-emotional development of disadvantaged children.

The possible harm of very early childcare due to the child's detachment from the parents may be caused rather by parental employment, i.e. maternal employment, than by exposure to institutional childcare itself. In a comparative study using longitudinal data from five OECD countries, Huerta et al. (2011) find evidence that, in some of the countries, maternal employment within the first six months after childbirth may have negative but small effects on child outcomes, in particular cognitive development. Yet, they suggest that institutional arrangements such as daycare might have a significant positive and persistent compensation effect on child development, in particular if it is of sufficient quality. And, despite the fact that maternal employment is growing, both parents tend to spend nowadays more developmentally relevant time with their children (see e.g. Bianchi, 2000).

3. Data description

The PRIMA cohort studies have been collected by ITS Nijmegen together with SCO-KO Amsterdam with a biannual frequency between 1994/95 and 2004/05 and studied by many social scientists, in particular to evaluate elementary school investments. The studies follow children at a large, nationally representative sample of elementary schools in the grades 2, 4, 6 and 8. About 30 children were randomly selected at each school for each of the studied grades. The nationally representative sample for the second grade of the PRIMA wave 2004/05 covers a total of 10,751 observations at 309 elementary schools (Driessen et al., 2006). Those children were born in 1998/99 and thus about six years old at the time of data collection. Of the nationally representative sample I use a sub-sample that covers full information in the set of variables that are relevant for my analysis, N=4616.

The PRIMA data originates from four sources: 1) teacher assessments of children's non-cognitive skills; 2) parental survey responses on family background and children's early childhood care and education; 3) schools registry information, e.g. on general parental and child characteristics; and 4) language and cognitive test scores from the school administrations for the testing period in the middle of the second school grades (March 2005).

I use three indicators of school readiness at the age of 6 for evaluating the effects of childcare attendance. Cognitive and Dutch language test scores come from the nationally comparable 'Cito' tests called *Ordenen* (= arithmetic test) and *Taal voor kleuters* (= Dutch language test for toddlers). Those tests are regularly taken at nearly all Dutch elementary schools. As a third indicator I aggregate teachers' assessments of children's non-cognitive performance in class in terms of work attitude, social behaviour, self-confidence and well-being to a single principal-component factor. For easier comparison I have standardized all three scores of school readiness to a mean of 100 and a standard deviation of 15 (see Appendix II.a, b and c).

Table II-1 gives an overview of the standardized test outcomes across groups of children with different childcare experiences, as described in the survey among parents.

Table II-1: Summary statistics on child outcomes by childcare arrangement

	Obs.	Mean	Std. Dev.	Min	Max
Dutch language test at middle of 2nd grade					
Only <i>home-based care</i> experienced	424	98.58	14.76	36.93	121.47
Some <i>centre-based care</i> attended	4192	102.29	13.58	34.89	123.44
Cognitive test at middle of 2nd grade					
Only <i>home-based care</i> experienced	424	99.94	14.23	52.06	125.26
Some <i>centre-based care</i> attended	4192	102.05	14.07	39.86	127.70
Non-cognitive assessment by teacher during 2nd grade					
Only <i>home-based care</i> experienced	424	100.57	14.31	46.41	135.79
Some <i>centre-based care</i> attended	4192	100.67	14.71	31.77	135.79

The sub-sample for which parents provided childcare information shows that about 91% of children experienced centre-based childcare. The survey provides sufficient information to identify which type of centre-based childcare, i.e. preschool kindergarten or daycare centre, the child attended as well as the duration of those childcare experiences. Three quarters of children attended a preschool kindergarten for some time, and nearly every second child attended a daycare centre. Despite the fact that only about one in ten children did not attend any childcare institution before going to school, the spells of centre-based childcare are rather short compared to the possible time children spend at home.

The centre-based childcare attendance can be described not only in terms of type of provision but also in terms of the attendance duration. Parents were asked whether their children attended a daycare centre or a preschool kindergarten and, if they did so, for how many years and half-days per week. I calculate the doses of half-days of childcare experiences by multiplying the reported years with the reported half-days per week,⁵ assuming that there are 40 weeks per year in which a centre is open. However, information about childcare attendance is not available for all children as not all parents have provided information on at least some of the underlying questions. For those who attended centre-based childcare, the average attendance spell was 206 half-day sessions in daycare centres and 129 half-day sessions in preschool kindergartens; a large share of children has attended both types of childcare.

4. Methodology

To estimate the effects of experiencing centre-based versus home-based childcare, I use a linear regression model estimated by Ordinary Least Squares (OLS) in which I correlate standardized outcomes of child i (CO_i) with an attendance dummy of centred-based childcare as treatment variable ($I_{1,i}$) and as conditioning set the duration and type of childcare ($I_{2,i}$), plus the usual control set of child (X_i) and family characteristics (F_i) that are used in such cross-sectional studies to evaluate effects of human capital investments on child outcomes and assumed to be exogenous. The family background factors presumably account for part of the child's development predispositions, to the extent that they are linked to inherited abilities. As there is no further information available on the quality of the provided intervention in the PRIMA dataset, the childcare evaluation is limited to the estimation of mean effects of centre-based childcare provisions. To get consistent standard error estimates, errors are clustered at the individual elementary school group to account for potential nesting of child outcomes ($N_{\text{cluster}}=654$).

$$CO_i = \beta_0 + \beta_1 X_i + \beta_2 F_i + \beta_3 I_{1,i} + \beta_4 I_{2,i} + \varepsilon_i \quad (II.1)$$

I control for the child's gender, age in months when taking the test⁶, and for belonging to an ethnic minority in terms of having at least one parent who was born abroad, and as family background factors I include a dummy variable for

⁵ A half-day session in the Netherlands, morning or afternoon, lasts on average about 5 hours at a daycare centre and 2.5-3 hours at a preschool kindergarten (see Eurydice, 2007; OECD Review Team, 1999; Van der Vegt et al., 2007).

⁶ At that time children have attended elementary school on average for about 23 months when being tested. The enrolment age is not available for about a third of observations; missing values are replaced by average values to keep the sample size up.

living in a single-parent household, the number of children in the household, and whether the mother has been employed for more than 12 hours per week, as well as a categorical indicator for the highest education level that the parents attained. The PRIMA dataset does not include information about household incomes; instead it provides a proxy indicator for the social milieu of the child's family background, which is a factor score. Moreover, I include the father's Dutch language skills and parental weekly reading of books, newspapers and magazines as proxy variables for the parents' cultural capital with respect to learning.⁷

The ultimate policy goal is to see not only whether any child-centred care experience is beneficial in addition to any home-based care, but also to see whether the type of childcare makes a difference and especially whether children from disadvantaged backgrounds benefit more than others. Therefore, I account for the duration of the childcare spell at daycare centres and preschool kindergartens separately, and I study moderation effects of childcare experiences on the equality of test outcomes across different groups of children. I am particularly interested in whether the impact of childcare experiences varies by gender as well as family background factors such as ethnic origin, parental education and single-parenthood.⁸ Those factors are frequently regarded as influencing the chances to do well in life and respectively with the risks of falling behind (see, e.g. Björklund & Salvanes, 2011; Bowles et al., 2005; Jencks, 1979; Leseman, 2002; Plug & Vijverberg, 2001; Wößmann, 2004). To test the heterogeneity of effects, I use estimation models with interaction terms.

4.1. Non-linear link between childcare spell and child outcomes

Despite the given evidence on the variation of childcare effects by duration and intensity (see section 2), most studies evaluating childcare investments use linear specifications for the relationship between childcare attendance and the measured child outcomes. However, effects may eventually be averaged out when using linear specifications. To address this issue with the correct functional form, I explore the relationship in the given data by using local regression smoothing (LOWESS) to trace out whether the relationship between the outcome variables

⁷ The Dutch language skills of mothers and fathers are highly correlated; I therefore include only the indicator for fathers' language skills. Other indicators of the parents' cultural capital such as the frequency of theatre and museum visits are not used as they are strongly correlated with parental reading.

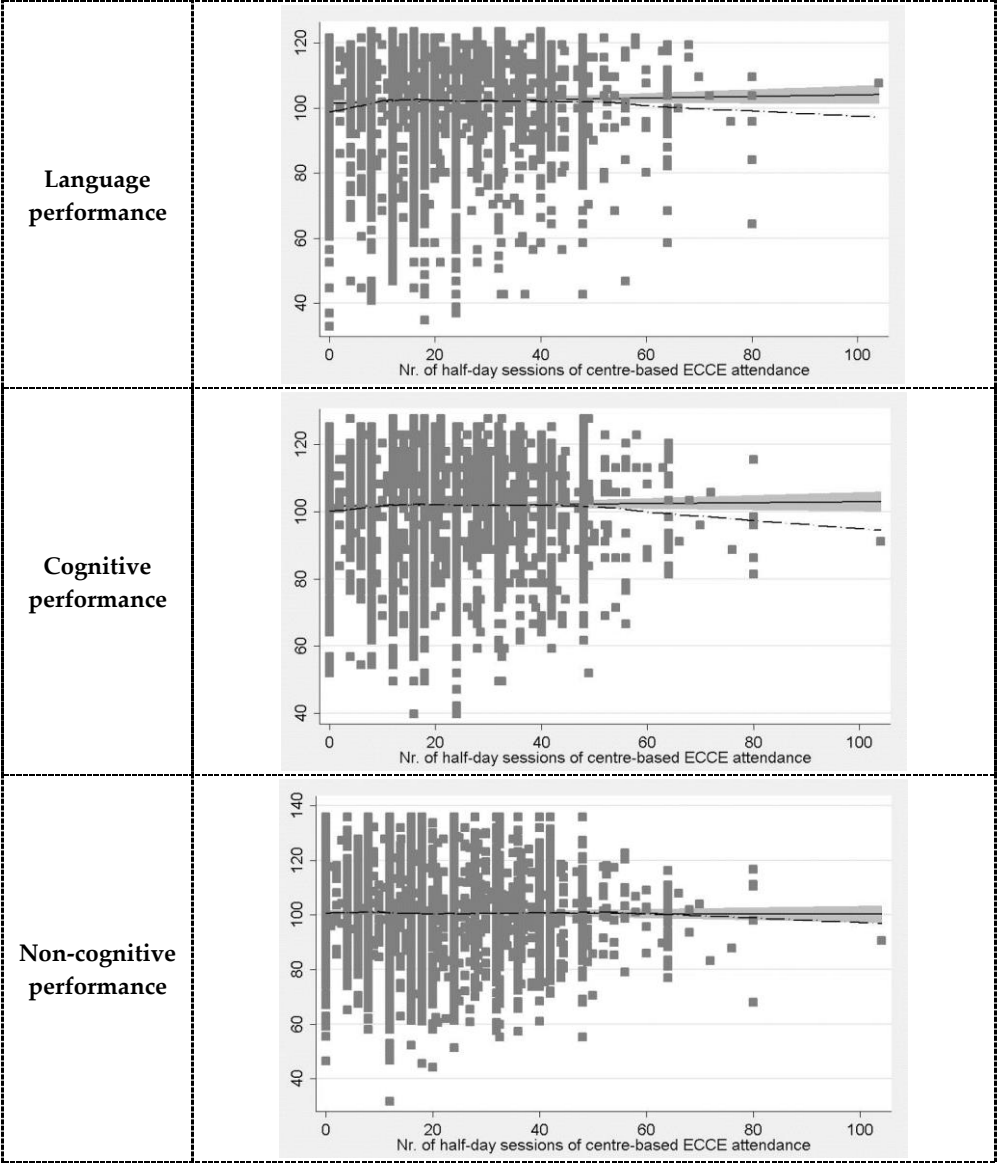
⁸ To keep the paper compact, I present only estimation output when I find significance heterogeneity in the treatment effects.

and the amount of childcare attendance has a non-linear pattern.⁹ Yet, it is important to note that the unconditional LOWESS graphs do not prove the existence of nonlinearity in the impact, since they may hide a potential compositional effect or sample sorting. If, for instance, parents were aware of a potential harm of too high intensity of preschool attendance it would be crucial whether they would have any choices with respect to using longer hours; i.e. the poor might more likely choose longer hours because they have to work, the rich might choose the optimal amount.

[For Figure II-1, please turn the page.]

⁹ Note that daycare centres are attended much earlier than preschool kindergartens. A too long daycare spell is likely to reflect a too early or too intensive enrolment of the child at the centre-based care. For preschool spells, non-linearity might reflect that too high intensity could be harmful.

Figure II-1: Link between child outcomes and centre-based childcare experience



Note: LOWESS smoothing (dotted line) of relationship between child outcomes (vertical axis) and childcare attendance (horizontal axis) and 95% confidence intervals (shaded area) around the fitted values (solid line).

These figures indicate non-linearity in the relationship between test scores and centre-based childcare attendance, in particular for language and cognitive skills. What could be the reason for such non-linear patterns? The most plausible argument for a levelling off in the relationship between childcare experience and test outcomes would be diminishing returns. This would imply that the marginal returns from additional days of centre-based childcare experiences decrease. In the framework of a human capital production function, this means that the inputs from preschool experience are not causing any additional learning effect for the child anymore, even if the child had stayed for a longer time in the preschool. The pattern indicates not only a levelling off but even negative marginal rates. A possible explanation for this may be that too intensive preschool attendance turns out to be harmful for the child.

Separate LOWESS estimates by different type of centre-based childcare indicate a stronger curvature for preschool kindergarten spells than daycare centre spells. Since the invention of the kindergarten by Friedrich Fröbel in the 19th century, the value and potential harm of early education has been discussed. In comparing the more structured, education-oriented approaches of preschools to daycare centres, critics regarded preschools as having too severe interference with children's needs for playful and unconstrained time to develop themselves in such institutions, causing in particular negative implications for children's non-cognitive development. Proponents have argued that more structured and targeted activities, as for instance to stimulate early math and reading preparedness, are beneficial for a child's development. If there are no hidden compositional effects, a non-linear relationship could reflect that both arguments may be valid.

Such two-dimensional response plots may provide a relatively easy first look at higher dimensions, outliers and other influential observations. Yet, the interpretation of individual frames in a scatterplot matrix is relatively easy only in isolation as a simple regression problem with a 1-to-1 relationship. But the interpretation of a scatterplot matrix can be less straightforward when viewed in the context of a full, multiple regression on a bigger conditioning set that may complicate matters (see Cook, 1998; chapter II). Hence, to interpret the pattern as a nonlinear relationship, I need to make some *a priori* assumptions regarding what I expect, namely that the function is likely to be smooth and single-peaked (see Leamer, 1983).

Schütz, Ursprung & Wößmann (2008) provide an alternative explanation: they suggest that if human capital accumulation at home is linear in time and if preschooling operates with decreasing returns, it could be possible that children acquire less human capital if they are sent to preschool for a too long time, as children may benefit too little from human capital accumulation at home. This argument may be linked to the fact that parents who send their child for more

hours per day to preschool also send the child more years to daycare centres. In this case children would spend insufficient time with their parents, which would manifest weaker test performance if parental influence has no diminishing returns. The data does not provide information on the ages when children attended either of the institutions, but there is some evidence that among children who have attended both types of childcare institutions (about 28%) the years spent in preschool are positively correlated to the years spent in daycare centres, while the enrolment age at elementary school is about the same for any subgroup of children.

To account for possible non-linearity in the relationship between childcare attendance and the child outcomes, I add a quadratic term of the approximated number of hours of childcare attendance in line with other researchers such as, for example, Landvoigt, Mühler & Pfeiffer (2007). Such quadratic terms have not been used in any of the major studies on Dutch early childhood education yet. As centre-based childcare spells can potentially be longer when a child enrolls in elementary school at a later age, I also control for the enrolment age (in months).

4.2. Instrumental variable (IV) application

When testing the cross-sectional regression models I cannot ignore possible joint determination of the dependent variable and the treatment variable, e.g. because important explanatory variables such as the unobserved initial ability endowment of the child before daycare attendance might be insufficiently accounted for by the included child and family background factors. This could cause the variation in childcare attendance to have unobserved (endogenous) heterogeneity and result in biased and inconsistent OLS estimates. This reflects a typical problem in evaluating the effects of human capital investments such as childcare and education when no controlled experiment is available or feasible. An alternative approach is to use an instrumental variable that does not determine the studied child outcome but provides information on exogenous variation on the potentially endogenous treatment allocation, which can be used for a two-stage least squares (2SLS) estimation (see, for example, Angrist, Imbens, & Rubin, 1996; Imbens & Angrist, 1994).

As potential sources of exogenous variation I use the aggregate daycare and preschool attendance at the larger geographic level of provinces; I find significant variation across the twelve provinces in preschool and daycare attendance, which reflects different progress between more urban and rural regions in extending centre-based childcare provisions (see Table II-7). A similar application of geographic aggregates as source of instrumental information on the treatment variable has been done by Dustmann & Preston (2001); they evaluated the impact of ethnic minority group concentration in the UK on the attitude towards those

ethnic groups. Heckman, Layne-Farra & Todd (1995) as well as Card & Krueger (1996) use aggregate statistics to evaluate the effect of school quality on earnings. They indicate the importance of non-random sorting of residence to identify consistent effect estimates. For that reason, I need to assume that parents do not choose their residence according to the available childcare provisions. Parents' influence on the overall preschool quality in the province is marginal, and it is unlikely that parents move across provinces due to daycare and preschool availability. Parental sorting within the province does not affect the overall composition in the province, and within-province sorting should therefore not affect validity of the IV estimation.

While the presence of such exogenous variation can only be argued theoretically, the strength of the correlation between the IV and the treatment variable can be confirmed by first-stage estimation results conditional on all the other explanatory variables. To assess the power of the IV application, a Wooldridge test (1995) can show whether the instrumented variables are exogenously determined and a Hausman test (1978) can assesses whether there are systematic differences between coefficients of IV and OLS estimates.

4.3. Structural equations modelling (SEM)

Instrumental variable applications formulate straightforward moment conditions without any explicit assumptions about a specific structural model on the selection into treatment. However, the chosen instrumental variable could turn out to be a weak predictor of the treatment variable. This can be tested, by looking at the first-stage correlation between the instrument and the studied treatment. Yet, the instrumental variable might provide only a local average treatment effect (LATE) estimate if the endogenous treatment effect is heterogeneous (Imbens & Wooldridge, 2007). In that case, approaching the issue of causality from a theoretical point of view by modelling the choice into treatment participation is an alternative strategy to get closer to unbiased effect inferences in a structural way.

Structural equations modelling (SEM) provides more flexibility than instrumental variable approaches by allowing correlated error terms and modifications to look at heterogeneous treatment effects, without additional assumptions, e.g. about linearity, as in IV applications (Bernal & Keane, 2010). It also allows a combination of various decisions in a structured way (see, e.g. Keane & Wolpin, 1997). And SEM provides more efficient parameter estimates than IV estimates as it considers all information simultaneously. However, SEM becomes unbiased only with large samples and possibly turns out to be more sensitive to violations of assumptions than IV estimates; errors in one strand of the system of relations that SEM estimates may have unknown effects throughout the system.

SEM is often used in the economic literature to guide empirical work or to make predictions as they explicitly allow for the construction of counterfactual results that are often lacking when evaluating treatment effects without having experimental data with a random treatment allocation at hand (Angrist & Krueger, 1999; Heckman & Macurdy, 1986). Despite this advantage, SEM examples are still rare in the childcare evaluation literature. As an exception, Arnold et al. (1998) apply 2SLS and SEM to evaluate the impact of teachers' laxness or over-reactive discipline in daycare on child behavioural outcomes. And in a more recent study Bernal and Keane (2010) use SEM to model maternal employment and childcare choices to 'quasi-structurally' evaluate the effect of non-maternal childcare for children of single mothers. They use variation in welfare rules across states as a plausible exclusion restriction, finding evidence for biases that are due to unobserved heterogeneity.

The system of equations of the SEM is assessed by a two-step estimation procedure, controlling directly for the correlation between the error term in the outcome equation with the treatment variable (Heckman, 1976). The binary treatment choice that is modelled in the first stage estimates at which threshold of explanatory variables the choice switches from mere early childhood care at home, i.e. by parents, relatives or nannies, to attending any centre-based childcare institution (see e.g. Heckman, 1978). The generated inverse Mills ratio augments the second stage estimation for the outcome equation to control for potential selection biases that are a result of omitted variables.

I model the choice into any childcare centre, under control for the child's background and the type and duration of childcare attendance. The choice for centre-based childcare is driven by a number of demand and supply factors. On the demand side, maternal employment is anticipated to be the most relevant factor, in particular among poorer or single parents. Since the late 1980s, the Dutch government has been promoting maternal employment through gradual extension of centre-based childcare provisions. Parents with more children are expected to choose centre-based childcare arrangements less often. Parental preferences as approximated by parental education, cultural capital and language skills are additional factors determining the demand for centre-based childcare. Parents with higher endowments of such factors are more likely to choose centre-based provisions, i.e. because they are better informed about how to access them. Choices are also determined by supply factors such as availability, user fees and quality attributes. Despite discounts in such fees for poorer parents, parents with higher incomes are more likely to be able to afford centre-based childcare (see e.g. Gathmann & Sass, 2012).

When assessing parental childcare choices, it is essential to understand whether parents perceive the quality of a specific centre-based childcare provision as being

higher or lower than home-based alternatives. Hence, in the selection equation, I also account for the type of centre-based care, a commonly used quality indicator. There might be trade-offs between different determinants of the decision for a specific childcare provision. Parents might prefer, for example, high-quality childcare arrangements if they choose to have fewer children (for a discussion of the quality-quantity trade-off, see e.g. Blau & Hagy, 1998; Ermisch, 1989). An extensive set of background factors is included in the choice models to account for them.

As an exclusion restriction, I use variation across provinces, degree of urbanisation at the parent's home, and whether parents adhere to a religious belief. These factors enter only the choice equation but not the outcome equation (see, e.g. Heckman, Lalonde, & Smith, 1999; Verbeek, 2000).¹⁰ Those three exclusion restrictions prove to be irrelevant in predicting the child outcomes but very important in predicting variation in centre-based childcare attendance.

In the first stage, I estimate a *probit* function of the treatment dummy for centred-based childcare attendance ($I_{1,i}$), on child characteristics (X_i), family background factors (F_i), the duration and type of childcare ($I_{2,i}$) and the restriction variables: indicators for each province (P_i), for the degree of urbanisation (U_i) and for having a religious faith. The estimated choice equation (2) is hence:

$$\Pr(I_{1,i}) = \gamma_0 + \gamma_1 X_i + \gamma_2 F_i + \gamma_3 I_{2,i} + \gamma_4 P_i + \gamma_5 U_i + \gamma_6 R_i + \delta_i \quad (II.2)$$

From these estimates, I obtain the inverse Mills ratio, which is the ratio of the probability density function of the predicted values to the cumulative distribution function of the predicted values. The inverse Mills ratio enters then in the second stage, the outcome equation (1).

5. Estimation results

I first assess the impact of centre-based childcare with OLS estimations and then turn to instrumental variable and structural equations model applications to address possible biases.

¹⁰ The Netherlands has 12 provinces. PRIMA sub-samples per province are not representative, yet there is no indication of any systematic sampling interference that could bias the relationship between childcare attendance and the province indicators. The urbanization degree ranges from 0 = "completely rural" to 5 = "completely urban". The indicator for religion reflects maternal religious beliefs, which are strongly correlated with paternal religious beliefs, 0.76. Among the mothers covered by the sample, about 35% do not adhere to any belief, 28% adhere to the roman-catholic church, 20% to reformed beliefs, 4% to other Christian confessions, 9% to the Islam, and less than 3% to other beliefs.

5.1. OLS estimates

I run a number of OLS estimation models for the three school readiness indicators, including a number of interactions with subgroup indicators – gender, parental education, single parenthood, social milieu of the household and ethnic origin. For convenience only estimation output of significant interactions is presented. To test the validity of non-linearity in the relationship of the treatment to the outcomes, I run the full estimation model with and without squared terms. I include both attendance spells together as they describe separate influences in early childhood.

[For Table II-2, see following page.]

Table II-2: OLS regression output

Estimated OLS Model	Dutch language skills			Cognitive skills			Non-cognitive skills	
	1	2	3	5	6	7	8	9
Childcare indicators								
Centre-based childcare (attendance dummy)	-0.92 [-0.88]		-1.14 [-1.10]	-1.26 [-1.23]		-0.3 [-0.26]	0.25 [0.22]	
Daycare attendance (in units of 10 half-day sessions)	0.05 [1.42]	0.02 [1.16]	0.05 [1.40]	0.07* [1.67]	0 [0.19]	0.07* [1.69]	-0.02 [-0.49]	-0.01 [-0.82]
Length of daycare attendance, squared term	0 [-0.14]		0 [-0.13]	0 [-1.02]		0 [-1.05]	0 [0.31]	
Preschool attendance (in units of 10 half-day sessions)	0.24*** [2.79]	-0.01 [-0.34]	0.24*** [2.79]	0.19** [2.09]	-0.05 [-1.50]	0.19** [2.07]	-0.04 [-0.41]	-0.09** [-2.56]
Length of preschool attendance, squared term	-0.01*** [-3.29]		-0.01*** [-3.30]	-0.01*** [-2.90]		-0.01*** [-2.85]	0 [-0.77]	
Subgroup interaction effects								
Centre-based childcare X low educated parents						-3.12* [-1.92]		
Centre-based childcare X single parent			9.49* [1.68]					
Child characteristics								
Gender (boy=1)	-2.40*** [-6.52]	-2.40*** [-6.46]	-2.40*** [-6.52]	-1.76*** [-4.53]	-1.75*** [-4.49]	-1.75*** [-4.49]	-5.76*** [-14.24]	-5.76*** [-14.28]
Age at time of testing (in years, middle of 2 nd grade)	8.36*** [14.99]	8.34*** [14.98]	8.43*** [15.07]	8.46*** [14.22]	8.43*** [14.17]	8.44*** [14.19]	1.63*** [2.59]	1.64*** [2.59]
Ethnic minority (min. one parent born abroad = 1)	-9.42*** [-7.76]	-9.86*** [-8.14]	-9.39*** [-7.78]	-4.25*** [-4.19]	-4.58*** [-4.52]	-4.35*** [-4.29]	5.22*** [5.28]	5.06*** [5.20]

(continued on next page)

Table II-2: OLS regression output (continued)

	Dutch language skills			Cognitive skills			Non-cognitive skills		
	1	2	3	5	6	7	8	9	
(continued)									
Estimated OLS Model									
Family background factors									
Parenthood (single-parents = 1)	-1.56 [-1.39]	-1.52 [-1.35]	-10.52* [-1.90]	-1.98* [-1.68]	-1.95* [-1.65]	-1.91 [-1.62]	-1.19 [-1.08]	-1.18 [-1.07]	
Number of children in household	-0.23 [-0.98]	-0.27 [-1.18]	-0.25 [-1.07]	-0.04 [-0.14]	-0.05 [-0.18]	-0.03 [-0.10]	0.83*** [3.33]	0.79*** [3.16]	
Mother's employment (>12hrs / week = 1)	-0.25 [-0.60]	-0.26 [-0.61]	-0.23 [-0.56]	-0.01 [-0.01]	-0.03 [-0.07]	-0.02 [-0.05]	0.6 [1.34]	0.61 [1.37]	
Highest parental education: level 2 (Middle / professional educ.)	3.83*** [5.51]	3.83*** [5.51]	3.86*** [5.56]	3.11*** [4.42]	3.08*** [4.38]	0.42 [0.27]	-1.28* [-1.78]	-1.26* [-1.77]	
Highest parental education: level 3 (Higher / professional / acad. educ.)	4.89*** [6.80]	4.85*** [6.79]	4.94*** [6.88]	5.06*** [6.85]	5.00*** [6.76]	2.35 [1.46]	-0.4 [-0.52]	-0.38 [-0.50]	
Parental reading (hours per week)	-0.05 [-1.46]	-0.05 [-1.45]	-0.05 [-1.48]	-0.03 [-1.02]	-0.03 [-1.03]	-0.03 [-0.97]	-0.07** [-1.97]	-0.07** [-1.98]	
Father's Dutch language skills	1.78*** [4.39]	1.80*** [4.44]	1.77*** [4.36]	0.62 [1.54]	0.64 [1.58]	0.62 [1.53]	-0.04 [-0.09]	-0.03 [-0.08]	
Social milieu (PRIMA VI factor score)	4.33*** [12.87]	4.32*** [12.87]	4.32*** [12.83]	4.76*** [13.64]	4.74*** [13.68]	4.77*** [13.66]	7.68*** [20.86]	7.69*** [20.91]	
Constant	26.21*** [6.52]	26.86*** [6.69]	26.15*** [6.52]	28.78*** [6.63]	29.18*** [6.65]	30.60*** [7.04]	63.75*** [13.35]	64.17*** [13.72]	
Observations	4616	4616	4616	4616	4616	4616	4616	4616	
Adjusted R ²	0.23	0.23	0.23	0.16	0.16	0.16	0.17	0.17	

Note: Test scores are standardized to mean 100, standard deviation 15; * significant at 10% level, ** at 5% level, and *** at 1% level; clustered at individual school group level (robust t-stat. in parentheses).

The included child and family background characteristics show the usually expected signs. Notwithstanding the rather large sample size, I do not find any significant correlation with the child outcomes for the indicators of maternal employment, and parental customary weekly reading is also only marginally significant.

Coefficients for daycare attendance are significant, which comes to a surprise given that the average daycare spells are longer than the average preschool spells. It may reflect that daycare attendance is neutral in comparison to parental childcare. As preschools have presumably been of higher quality than daycare centres in the early 2000s, I would expect to find some significant relationships, at least for the preschool attendance indicators. None of the centre-based childcare estimates are significant. There is only a significant mean effect of the preschool attendance for non-cognitive skills, but it points to a negative relationship. This may reconfirm the argument that overly structured childcare experiences could turn out to be harmful. However, if I account for possible non-linearity I find a significant positive main effect of preschool attendance on language and cognitive outcomes, with decreasing returns to more attended half-days, although the effects are very small in scale.

Frequently, focusing on the mean as a measure to identify effects (as done with OLS) results in a disregard of properties at the end of the tails of the distribution or of distorting impacts of outliers. The estimated effects of centre-based childcare might be sensitive to such distributional aspects – effects might be more significant at other quantiles than at the mean or even show opposite signs (see Koenker & Bassett, 1978). Therefore, I estimate the impact of centre-based childcare attendance also on different quantiles of the conditional distribution of the error term (see Table II-8 in Appendix II).

Quantile regressions reconfirm largely that centre-based childcare does not significantly predict child outcomes. However, effect estimates by decile indicate some evidence for negative effects of centre-based childcare at the lowest end of the distribution for language and cognitive development; this could suggest that centre-based childcare could be harmful for disadvantaged children. There is some evidence for positive effects of centre-based childcare on child language and non-cognitive outcomes closer to the median.

When looking at the variation of childcare attendance effects across various subgroups whose backgrounds are presumably disadvantageous, I cannot find any effect of heterogeneity for the majority of such subgroups. However, there are significant correlations for some of the interactions with the group of children with less educated parents and children of single parents; centre-based childcare shows a more positive effect for children of single parents, but a rather negative effect for children of less educated parents.

When running the regressions by potentially disadvantaged subgroup and by the main group separately, neither output shows any significant effects from centre-based childcare. There is one exception, though. For the sub-sample of children from ethnic minorities, centre-based childcare shows marginally significant effects on language development and strongly significant effects on cognitive development – both effects being negative.

5.2. IV estimates

In the following, I try to assess whether omitted variable bias may have influenced the OLS results. I use the aggregated mean of the child-centred childcare dummy at the province levels to instrument for the choice for centre-based childcare.

Table II-3: 2SLS regression output, applying the instrumental variable

a. 1st stage results

Dependent variable	Centred-based childcare (attendance dummy)
Instrumental variable:	
Province aggregate of child-centred childcare	0.46*** [9.44]
1. Joint significance of the instrument(s) in the first stage; critical values $F > 10$	$F = 41.13$ $p = 0.00$ IVs strong

Note: Robust t-statistics in parentheses; * significant at 10% level ($|t| > 1.64$), ** at 5% level ($|t| > 1.96$); *** at 1% level ($|t| > 2.58$). A critical F value for the test on joint significance of the IVs (see test 1) is 10 (as stated, for example, in Baum, Schaffer, & Stillman, 2003).

The first-stage output of the 2SLS-regressions shows that the province aggregates strongly predict the individual spells (see Table II-3.a).

b. 2nd stage results

	Dutch test (M2)		Cognitive test (M2)		Non-cognitive assessment	
Estimated model	OLS (2)	2SLS (1)	OLS (6)	2SLS (2)	OLS (9)	2SLS (3)
Childcare indicator						
Centred-based childcare (attendance dummy)	-0.92 [-0.88]	-10.04 [-1.07]	-1.26 [-1.23]	-18.66* [-1.76]	0.25 [0.22]	13.86 [1.21]
Attendance variables	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES
Family background factors	YES	YES	YES	YES	YES	YES
Constant	26.21*** [6.52]	30.29*** [5.20]	28.78*** [6.63]	36.56*** [6.16]	63.75*** [13.35]	57.66*** [8.55]
Observations	4616	4616	4616	4616	4616	4616
Adjusted R ²	0.23	0.21	0.16	0.10	0.17	0.14

Note: Robust z-statistics in parentheses, clustered at individual school group level (=654 clusters); * significant at 10% level ($|z|>1.64$), ** at 5% level ($|z|>1.96$); *** at 1% level ($|z|>2.58$).

The 2nd stage output of the 2SLS regressions shows no significant effects for child-centred childcare attendance on language and non-cognitive outcomes. However, child-centred childcare attendance is marginally significant in determining cognitive outcomes.

c. Post-estimation model tests

Dependent variable	Dutch test (M2)	Cognitive test (M2)	Non-cognitive assessment
Estimated 2SLS model	2SLS (1)	2SLS (2)	2SLS (3)
2. Wooldridge's test of exogeneity of instrumented variables	F=1.00 p=0.32 endogenous	F=3.16 p=0.08 (marginally) endogenous	F=1.44 p=0.23 endogenous
3. Hausman test for systematic differences between 2SLS and OLS estimation model	$\chi^2=13.16$ P=0.44 No diff.	$\chi^2=13.06$ P=0.44 No diff.	$\chi^2=4.46$ P=0.99 No diff.
Interpretation of IV results	Prefer OLS	Prefer OLS	Prefer OLS

Note: I test for each of the two stage least squares models whether it is better than an ordinary least squares specification, i.e. I use a Wooldridge test (1995) to see whether OLS are providing consistent estimates because the instrumented variables are actually exogenously determined. And finally I use a Hausman test (1978) to check for systematic differences in the consistency and efficiency in the specifications of each IV and OLS estimation model.

Post-estimation tests show that there is reason to worry about endogeneity (see Wooldridge's test of exogeneity) and thus to look for a better identification strategy than OLS estimates. However, the 2SLS estimates result in no systematic difference to the OLS estimates. This indicates that the chosen IV might be too weak or heterogeneous in predicting the treatment variable; an alternative estimation strategy may be needed.

5.3. SEM estimates

Structurally estimating the choice for the mode of childcare provides an alternative method to amend possible selection bias of effect estimates. The following table shows the *probit* estimates of the childcare choice model.

Table II-4: SEM 1st Stage - *probit* centre- versus home-based childcare choice

Dependent variable	Centred-based childcare (<i>attendance dummy</i>)
Child characteristics	
Gender (<i>boy=1</i>)	0.09 [1.49]
Age at time of testing (<i>in years, middle of 2nd grade</i>)	0.00 [-0.00]
Ethnic minority (<i>min. one parent born abroad = 1</i>)	0.05 [0.39]
Family background factors	
Parenthood (<i>single-parents = 1</i>)	0.17 [0.93]
Number of children in household	-0.47*** [-15.42]
Mother's employment (<i>>12hrs / week = 1</i>)	0.22*** [2.76]
Highest parental education: level 2 (<i>Middle / professional educ.</i>)	0.34*** [4.07]
Highest parental education: level 3 (<i>Higher / professional / acad. educ.</i>)	0.47*** [4.99]
Parental reading of books, newspapers or magazines (<i>hours per week</i>)	-0.01*** [-3.12]
Father's Dutch language skills	0.02 [0.41]
Social milieu (<i>PRIMA VI factor score</i>)	0.17*** [3.62]

Exclusion restrictions	
Drenthe	0.08 [0.39]
Flevoland	-0.43*** [-2.68]
Friesland	0.00 [0.01]
Gelderland	-0.17 [-1.58]
Groningen	-0.62*** [-3.84]
Limburg	0.17 [1.18]
Noord-Brabant	0.51*** [3.84]
Noord-Holland	0.25** [2.16]
Overijssel	0.54*** [2.77]
Utrecht	-0.67*** [-3.88]
Zeeland	-0.44*** [-2.95]
Zuid-Holland	<i>dropped</i>
Degree of urbanisation (completely rural = 1 / completely urban = 5)	0.08** [2.51]
Religion of the mother (no religious affiliation = 1)	0.53*** [6.46]
Constant	1.26** [2.07]
Observations	4607
Adjusted R ²	0.26

Note: The sub-samples of schools are not representative for the provinces.

The probit estimates for the first step of the structural model show that child characteristics are not determining the choice for centre-based care, whereas family background factors show the expected significance and directions. Parents with more children tend to send their children less often to a centre-based care arrangement. Better educated parents, those from higher social milieus, and employed mothers tend to opt more often for centre-based childcare.

The chosen exclusion restrictions – province indicators, degree of urbanisation and mothers' religiosity - are significant determinants of the choice for centre-based childcare. In some provinces, such as Noord-Brabant, Noord-Holland and Overijssel, parents are more likely to send their child to centre-based care arrangements; in other provinces, i.e. Flevoland, Groningen, Utrecht and Zeeland,

parents are less likely to do so. In more urban areas, it is more likely that parents will chose a centre-based care arrangement. More religious parents tend to keep their children at home.

Table II-5: SEM 2nd Stage - OLS regression model including Mills ratio

Dependent variable	Dutch language skills	Cognitive skills	Non-cognitive skills
Estimated SEM Model (2 nd stage)	1	2	3
Childcare indicators			
Centred-based childcare (attendance dummy)	-0.70 [-0.73]	-0.60 [-0.59]	-0.18 [-0.17]
Daycare attendance (in units of 10 half-day sessions)	0.05 [1.33]	0.06 [1.56]	-0.02 [-0.43]
Length of daycare attendance, squared term	0.00 [-0.12]	0.00 [-0.95]	0.00 [0.24]
Preschool attendance (in units of 10 half-day sessions)	0.24*** [2.77]	0.19** [2.04]	-0.04 [-0.42]
Length of preschool attendance, squared term	-0.01*** [-3.37]	-0.01*** [-2.87]	0.00 [-0.72]
Child characteristics			
Gender (boy=1)	-2.40*** [-6.70]	-1.72*** [-4.49]	-5.78*** [-14.55]
Age at time of testing (in years, middle of 2 nd grade)	8.34*** [16.33]	8.44*** [15.43]	1.59*** [2.81]
Ethnic minority (min. one parent born abroad = 1)	-9.53*** [-12.58]	-4.48*** [-5.52]	5.42*** [6.44]
Family background factors			
Parenthood (single-parents = 1)	-1.56 [-1.59]	-1.96* [-1.86]	-1.19 [-1.09]
Number of children in household	-0.39 [-1.24]	-0.61* [-1.82]	1.21*** [3.48]
Mother's employment (>12hrs / week = 1)	-0.20 [-0.48]	0.16 [0.37]	0.47 [1.04]
Highest parental education: level 2 (Middle / professional educ.)	3.91*** [7.04]	3.43*** [5.76]	-1.42** [-2.31]
Highest parental education: level 3 (Higher / professional / acad. educ.)	5.02*** [8.03]	5.51*** [8.23]	-0.64 [-0.92]
Parental reading of books, newspapers or magazines (hours per week)	-0.05* [-1.76]	-0.05 [-1.56]	-0.06* [-1.73]
Father's Dutch language skills	1.81*** [5.21]	0.72* [1.92]	-0.10 [-0.25]
Social milieu (PRIMA VI factor score)	4.37*** [15.37]	4.89*** [16.06]	7.60*** [24.12]

Control function <i>(Mills ratio)</i>	1.08 [0.68]	3.92** [2.31]	-2.57 [-1.46]
Constant	26.10*** [7.03]	27.94*** [7.03]	64.53*** [15.68]
Observations	4607	4607	4607
Adjusted R ²	0.23	0.16	0.17

Note: Test scores are standardized to mean 100, standard deviation 15; * significant at 10% level, ** at 5% level, and *** at 1% level; clustered at individual school group level (robust t-stat. in parentheses).

The control function estimates once again reconfirm that there is no significance of coefficients on centre-based childcare attendance. Given that the tested structural equations reflect the true character of selection into centre-based childcare, the partial insignificance of the coefficient on lambda (the inverse Mill's ratio) suggests that the selection bias is not very large in this model. And indeed, the coefficient estimate on the treatment variable changes little compared to the simple OLS model Table II-2. For cognitive skills, the significance of lambda suggests that the control function works better, indicating some selection bias. Yet, the coefficient of the treatment variable is not significant, which is in line with the OLS estimates.

6. Conclusions

The last decades have seen a substantial increase in maternal employment in the first year after giving birth. When skill gains multiply over a lifetime (see, e.g., Heckman, 2008), small initial skill returns from centre-based early childhood care and education can turn out to produce larger long-term effects.

My analysis shows that previous studies may have prematurely concluded that there are indeed significant short-term effects of centre-based childcare attendance on child outcomes. Attempts to deal with potential omitted variable biases by applying instrumental variables and structural equation models reconfirm ordinary least square estimates.

The results are based on different types of tests to assess the impact of centre-based childcare. Ordinary least square estimates assess the average treatment effect of centre-based childcare on school readiness. Instrumental variable estimates provide local average treatment effects that are corrected for possible biases, local implying that the instrument may not be valid across the whole range of treatment. And finally, I use structural equations modelling to account for possible biases. All tests point in the same direction, namely an absence of significant results. The structural estimation seems to work better with cognitive skills; this indicates that, if at all, cognitive skills seem to be the most affected, which is entirely consistent with the role of childcare.

The absence of significant effect estimates contradicts the evidence provided by Van de Vegt, Studulski & Klopogge (2007), who find some indication of positive effects of preschooling on language and cognitive outcomes, as well as by Magnuson, Ruhm & Waldfogel (2007), who find a negative main effect of pre-kindergarten attendance on non-cognitive short-term child outcomes. They reconfirm the findings of Driessen (2004; see also Driessen & Doesborgh, 2003), who studied the effectiveness of centre-based childcare for an earlier period using co-variation techniques and stressing that controlling for background factors leads to a disappearance of significant effect estimates.

When taking a closer look at how centre-based childcare experiences affect children who are at risk of falling behind in their later schooling, my results do not confirm the expected positive impacts or even point towards weaker cognitive outcomes for children of less educated parents. Only children of single parents are associated with higher language scores. Quantile regressions indicate that centre-based childcare may be harmful for disadvantaged children that are more likely to be represented in the lower end of the distribution for language and cognitive outcomes but beneficial for children with outcomes closer to the median with respect to language and non-cognitive development.

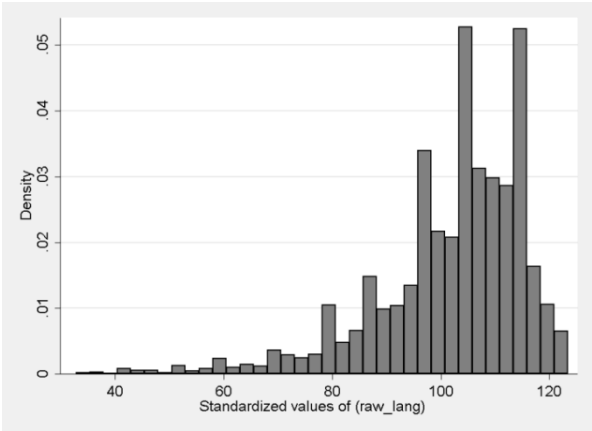
The lack of evidence on mean effects of centre-based childcare on all child development domains, combined with the indicated harmful effects for more disadvantaged children, raises questions about the quality and length of childcare spells that children of different background have actually experienced. Do disadvantaged children fare better if they attend higher quality centre-based childcare or attend shorter spells? I do find some evidence that there is a non-linear relationship between spells of centre-based childcare and child outcomes, in particular between preschool kindergarten attendance and language and cognitive development. By accounting for non-linearity in the spell of preschool kindergarten and centre-based daycare respectively, I partially control for potentially averaging out effects.

I conclude, in line with the literature, that the effects of centre-based childcare may be averaged across all qualities of daycare centres and preschools, preschooling being more likely to stimulate child development, and daycare attendance not being harmful. Further research is needed to identify whether there is stronger evidence for the effect of quality differences on childcare experiences.

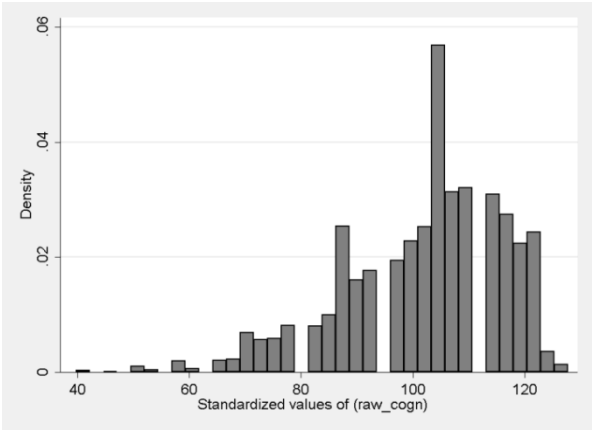
Appendix II

Figure II-2: Distribution of standardized child outcomes

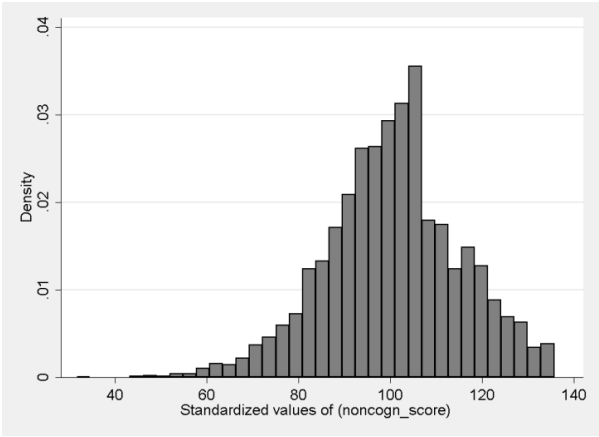
a. Dutch language test, M2



b. Cognitive test, M2



c. Non-cognitive assessment by teacher



Note: Cito test scores have been standardized to a mean of 100 and standard deviation of 15.

Table II-6: Descriptive statistics

Control variable	N	Mean	Std. Dev.	Min.	Max.
Dutch test (<i>taal voor kleuters, M2-testing moment</i>)	4616	101.945	13.731	32.9	123.4
Cognitive test (<i>ordenen, M2-testing moment</i>)	4616	101.852	14.093	39.9	127.7
Non-cognitive assessment score (<i>teacher's assessment at moment of survey</i>)	4616	100.661	14.672	31.8	135.8
Daycare attendance (<i>in units of 10 half-day sessions</i>)	4616	8.395	13.309	0.0	80.0
Preschool attendance (<i>in units of 10 half-day sessions</i>)	4616	9.706	6.991	0.0	40.0
Gender (<i>boy=1</i>)	4616	0.511	0.500	0.0	1.0
Age at time of testing (<i>in years, middle of 2nd grade</i>)	4616	5.983	0.352	5.3	7.6
Ethnic minority (<i>min. one parent born abroad = 1</i>)	4616	0.084	0.277	0.0	1.0
Parenthood (<i>single-parents = 1</i>)	4616	0.035	0.183	0.0	1.0
Number of children in family	4616	2.476	0.915	1.0	5.0
Mother's employment (<i>>12hrs / week = 1</i>)	4616	0.303	0.459	0.0	1.0
Highest parental education: level 1 (<i>no / elem. / lower educ.</i>)	4616	0.175	0.380	0.0	1.0
Highest parental education: level 2 (<i>middle / professional educ.</i>)	4616	0.448	0.497	0.0	1.0
Highest parental education: level 3 (<i>higher / professional / acad. educ.</i>)	4616	0.377	0.485	0.0	1.0
Parental reading (<i>of books, newspapers or magazines</i>) (<i>hours per week</i>)	4616	11.164	6.339	0.0	36.0
Father's Dutch language skills	4616	4.579	0.602	1.0	5.0
Social milieu (<i>PRIMA VI factor score</i>)	4616	3.871	0.702	1.0	5.0
Degree of urbanisation (<i>completely rural = 1 / completely urban = 5</i>)	4614	3.410	1.214	1.0	5.0
Religion of the mother (<i>no religious affiliation = 1</i>)	4614	0.356	0.479	0.0	1.0

Table II-7: Variation of average childcare attendance at each province

Province	Children in elementary schools			Attended preschool half-days		Attended daycare half-days	
	total population 2004/05	PRIMA 2004/05 sample	Sample coverage	Mean	Std. Dev.	Mean	Std. Dev.
Drenthe	5,807	142	2.45%	9.577	4.950	3.039	8.180
Flevoland	5,342	126	2.36%	9.683	9.013	6.991	11.847
Friesland	7,768	310	3.99%	10.317	5.441	4.438	10.118
Gelderland	24,264	748	3.08%	9.040	6.639	7.565	12.499
Groningen	6,280	154	2.45%	9.377	6.688	3.952	10.346
Limburg	12,007	357	2.97%	10.793	8.099	10.930	13.656
Noord-Brabant	28,564	623	2.18%	11.148	6.749	8.790	12.311
Noord-Holland	29,834	890	2.98%	9.418	6.780	10.425	15.319
Overijssel	13,800	225	1.63%	10.133	5.616	5.868	11.091
Utrecht	14,323	104	0.73%	8.068	7.335	7.873	12.701
Zeeland	40,190	196	0.49%	9.398	7.200	7.702	11.101
Zuid-Holland	4,527	741	16.37%	9.011	7.784	10.109	14.966
Total	192,706	4,616	3.47%	9.706	6.991	8.395	13.309

Note: The sub-samples of schools are not representative for the provinces. Figures on the total number of children aged 6 in 2004/05 in elementary schools are taken from CBS-Statline.

Table II-8: Simultaneous quantile regression

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Dutch skills	-3.91*	-0.63	-0.75	1.12*	0.32	-0.23	0.20	0.56	1.05	-5.01
	[-1.90]	[-0.49]	[-0.73]	[1.82]	[0.31]	[-0.23]	[0.22]	[0.59]	[1.48]	[-0.55]
Cogn. skills	-0.83	-3.25*	-1.78	-1.11	-1.20	-0.87	-0.85	0.33	-0.17	-7.61
	[-0.37]	[-1.79]	[-1.17]	[-0.97]	[-1.00]	[-0.61]	[-0.65]	[0.33]	[-0.15]	[-1.18]
Non-cogn. skills	-2.48	0.20	0.17	0.23	0.92	2.13*	1.29	0.64	0.15	-6.24
	[-1.44]	[0.14]	[0.12]	[0.17]	[0.60]	[1.76]	[0.75]	[0.38]	[0.07]	[-1.20]

Note: Effect estimates for centre-based childcare when running the OLS model (1) by decile, including childcare attendance indicators, child characteristics and family background factors. Test scores are standardized to mean 100, standard deviation 15; level t-statistics based on bootstrapped errors in brackets; * significant at 10% level, ** at 5% level, and *** at 1%.

Chapter III - Potential gains from child-centred early childhood education (*study II/III*)

1. Introduction¹¹

“Investing in disadvantaged young children is a rare public policy with no equity-efficiency trade-off. It reduces the inequality associated with the accident of birth and at the same time raises the productivity of society at large.” This argument of Heckman & Masterov (2007) summarizes the economic rationale of investing in early childhood education so as to improve the starting position of children with underprivileged family backgrounds. Early childhood care provisions are extended in many countries, predominantly to facilitate maternal employment; their coverage grew from 60.4% in 2000 to 62.8% in 2008 (OECD, 2011).

Introducing elements of Early Childhood Education (*generally abbreviated as ECE*), in addition to mere childcare services, receives growing political currency. Policymakers in the OECD countries see in extensions of ECE programmes a strategy to deal with over-aging and shrinking populations by supporting more children in achieving their full skills potential. Similarly, policymakers in less developed countries consider ECE programmes to be tools that empower more children to contribute to the growth of their knowledge-based economies and prepare them for a growing demand for higher skills in the labour market. However, when the state takes over part of the job to prepare children for their later lives, it is essential that it ensure sufficient quality for its childcare services.

The market for early childhood education programmes is growing in most countries. UNICEF (2008) reports that 15 of 25 studied OECD countries already provide subsidized and accredited early education services for at least 80% of 4-year-olds. The Netherlands, on which this chapter focuses, is not yet among them. Despite extending early childhood care and education provisions, policymakers are still insufficiently informed about the effectiveness of particular ECE approaches; the empirical evidence is rather limited. Particularly little is known yet about whether ECE programmes should stimulate child development and school readiness by providing a structured, curriculum and teacher-based care approach with direct instruction or rather a more flexible, child-centred approach where the ‘teacher’ guides the child based on individual needs, usually along a specific theme.

¹¹ An earlier version of this chapter has been released as UNU-MERIT Working Paper 2012(16).

Both types of programmes are frequently applied; however, most of the OECD countries, including the Netherlands, apply more ECE programmes with a teacher-based approach; Nordic and some central European countries as well as New Zealand follow more child-centred approaches (Nap-Kolhoff, Schilt-Mol et al., 2008; OECD, 2007). Parental preferences for a specific type of ECE provision are determined by their socio-economic but also cultural background. For example, Asian parents generally tend to prefer more teacher-based ECE programmes; Western parents show more preference for child-centred provisions (see, e.g. Tobin, Wu, & Davidson, 1991).

Constructivist theories of early child development, which go back to theorists like Piaget, Montessori & Vygotsky, stress the need for children to construct their own learning process and to build their own relationships with peers and teachers; whereas teacher-based preschool programmes, based on environmental theorists such as Skinner, Watson and Bandura, emphasize the importance of curriculum-based knowledge transmission (see, for example, Hohmann, Banet, & Weikart, 1979). The discussion about types of ECE is not always fully exclusive in terms of which approach to choose. Copple and Bredekamp (1997; chapters 4 and 5) stress that teachers' training and behaviour should accommodate a *Developmentally Appropriate Practice* via a wide range of teaching strategies, including well-adjusted curriculum-based guidance as well as individualized development stimulation. Evaluating the attendance of Chicago Child-Parent Centres in the mid-1980s, Graue et al. (2004) highlight the importance of an appropriate balance between teacher-direct instructional approaches, child-initiated activities and parental involvement to generate any long-term effects.

ECE programmes combine a set of didactic principles, learning approaches, educational materials and staff training to assure sufficient quality of centre-based care. Nonetheless, there is surprisingly scarce empirical evidence for the many questions that can be raised regarding the appropriate quality of childcare. A crucial question in this debate is whether children actually benefit from an educational stimulus in their early childhood and, if they do so, whether an ECE programme is more stimulating if it uses a programme-directed or rather child-centred approach.

Traditional beliefs that are more prevalent in Anglo-Saxon countries and France consider early childhood education as a means to socialise children - to make them ready for school. Instead, modern social pedagogic beliefs, which are more prevalent in Nordic and Central European countries, consider ECE to be supportive in the development of individual talents and stimulation of the families (Leseman, 2002; OECD, 2006). The share of child-centred ECE applications in OECD countries is increasing; a well-known example of such child-centred approaches is the Italian *Reggio Emilia* method. The Netherlands has recently

accredited *Speelplezier*, a newly developed child-centred ECE programme, as being in theory effective in raising the readiness for elementary schooling of children (see Pennings, 2009).¹² Accreditation of an ECE programme allows preschools to receive public funding for employing specially trained staff.

Before receiving official accreditation, the ECE programme *Speelplezier* has been used since the school-year 2000/01 in some of the preschools and schools in a few bigger southern municipalities in the Dutch province of Limburg. Rich data has been collected at the initial stage of all elementary schools in this region, allowing new effect evaluations of preschool kindergarten experiences. The theoretically proven effectiveness of child-centred approaches can thus be tested by a comparison of *Speelplezier*, and the few other child-centred ECE applications, with other alternative preschool approaches. Child-centred ECE approaches are ideally compared versus the two alternatives, preschooling with a teacher-based approach or preschooling without any early education at all. Such comparisons contribute valuable empirical evidence to the scarce Dutch and international literature and go beyond the theoretical argumentation about effectiveness that is still predominant in the policy arena.

The introduction of the novel Dutch ECE programme *Speelplezier* and the few alternative child-centred ECE approaches used in some municipalities of South Limburg is a unique opportunity to provide new empirical evidence. Despite annual aggregate monitoring reports of municipalities using child-centred ECE approaches, there is no empirical evidence on the effectiveness of those approaches yet. For that reason I address the following question in this chapter: *Do pupils perform better at school entry if they had attended a preschool using a child-centred ECE programme rather than any alternative preschool, which applies a teacher-based approach or no ECE approach at all?*¹³ *And does child-centred ECE benefit disadvantaged children above average?*

The presented empirical results are the first pupil-level evaluation of the ECE programme *Speelplezier* and its potential to reduce inequalities in educational

¹² The name *Speelplezier* means “fun in playing,” which reflects the programme’s uniqueness in trying to reach the desired development gains through its extensive attention to play activities that are, to a large extent, freely determined by the children themselves and in part guided by childcare and teaching staff. At least one staff member for each 25 children receives training over two to three years in using the education materials with themes to play and interact with the children and within the four different programme modules (see Wouterse-Schmitz, 2006).

¹³ I compare child-centred ECE to the pool of alternative preschool experiences, i.e. those with teacher-based ECE and those with no ECE applications. The sub-sample of teacher-based applications is too small for a direct comparison.

attainments. *Speelplezier* is the predominant child-centred ECE approach used in the studied region of South Limburg; more than 90% of children with child-centred ECE experiences attended a *Speelplezier* preschool.¹⁴

Several municipalities in South Limburg are currently expanding their public ECE investments, piloting a future development in the rest of the country. Such expansions should be guided by evidence on the effectiveness of ECE approaches, which in the studied case of South Limburg is primarily *Speelplezier*. Assessing the application of child-centred ECE in South Limburg provides evidence whether governmental scaling up of early childhood education investments is likely to deliver the expected positive results, in particular for the targeted children with disadvantaged backgrounds.

A description of ECE policies in the Netherlands (Section 2) precedes a review of the international and Dutch evaluation literature on early childhood education (Section 3). Then a description of the data (Section 4) and the empirical strategy (Section 5) paves the way for the presentation and discussion of the estimation results (Section 6). The chapter concludes with policy recommendations (Section 7).

2. Early childhood education policies in the Netherlands

In the Netherlands, a growing share of women is active in the labour market - 65.5% in 2000 and 73.5% in 2008. Many are in part-time contracts – 59.8% in 2008; more than double the OECD average of 25.3% (OECD, 2011). Whereas daycare centres or nursery (*kinderdagverblijven*) are the most extensive provisions of centre-based childcare, preschool kindergartens (*peuterspeelzalen*) cover a universal group of children with different backgrounds for a few half-day sessions (mornings or afternoons) each week.

Predominant centre-based childcare arrangements are daycare centres, which provide care for children from a few months after birth and later include out-of-school arrangements. Daycare provisions are largely organised by (semi) private organisations and have on average a lower quality standard than preschool kindergartens, largely because the latter started earlier in focusing on child development aspects and promoting professionalism.¹⁵ In the Netherlands, ECE

¹⁴ Other child-centred ECE approaches used in the studied region are, for example, *Startblokken* and *Basisonwikkeling*, but *Speelplezier* is applied by the vast majority of preschools using child-centred ECE.

¹⁵ Recently, the Dutch government promotes access to daycare centres and quality of such daycare provisions; however, for the studied children, provisions have been in general of lower quality than preschool kindergarten provisions.

programmes are applied in preschool kindergartens that provide care for children aged 2-4 years and at the initial stage of elementary schools, which are attended by children aged 4-6 years. Daycare centres are only starting to invest into ECE approaches.

Dutch ECE programmes are meant to stimulate linguistic, cognitive and social development. Policies to provide extra stimulation for disadvantaged children were in the 1980s predominantly focused on parent-oriented investments. The growth in female employment was accompanied by increasing demand for non-parental care. Initially the state did not respond directly to this demand; instead, charity foundations started to fill the gap and founded numerous early childhood care and education arrangements (Pfau-Effinger, 1999). The first preschools were set up in the 1970s (Nap-Kolhoff & Schilt-Mol, 2008). Since the 1990s centre-based childcare investments combined with parent-oriented programme elements have become a preferred public policy strategy.

In order to provide extra stimulus for disadvantaged children, the Dutch government started providing subsidies and regulating the quality of the childcare sector. Nowadays, the government wants to advance the development of early childhood education programmes by making them more professional in accordance with what (international) research identifies as most effective practices. And the Dutch government has announced that programmes will be extended to reach more disadvantaged children in need of such development stimulus (see, for example, OCW, 2008).

Evidence suggests that poverty, low social class, low educational attainment levels, to some extent traditional child-rearing beliefs and cultural background factors, foreign mother-tongues and low literacy can all have pervasive influences on informal education at home and hence are likely to result in insufficient preparation for school (see, e.g. Leseman, 2002). If, for instance, neither of the parents is using Dutch in conversations with their child, the child is more likely to face difficulties when entering the school. Since the 1990s, language stimulation has been the primary target of such ECE programmes, acknowledging particularly the problems of integrating children with migrant backgrounds into the Dutch elementary schools. Cognitive development has been a wider goal with respect to generally targeting ECE programmes towards children of less educated parents; since 2006, programme targeting related to migrant backgrounds has been phased out. Targeting the age group of 2.5 to 6 years acknowledges that this period is deemed sensitive for children's cognitive development and for their language advancements (Cunha et al., 2006; Heckman, 2008).

Starting in the year 2000, the first investments in ECE programmes were made in the four largest Dutch municipalities as they have bigger migrant communities. Since August 2002, ECE regulations became part of the *Municipal Compensatory*

Education Policy (see GOA § 6.2.3., OCW, 2001). At that time 27 mid-sized municipalities also received public investments for ECE programmes - in the studied region of South Limburg, these were Heerlen, Sittard-Geleen and Maastricht.

At the moment, there are efforts underway to extend ECE programmes in particular to smaller municipalities in the countryside as the focus of targeting is widened to all children whose parents are less educated. A significant number of those children are disadvantaged native children who live in rural areas. Major examples of these extension policies are governmental pilot projects carried out in the rural regions of South Limburg and East Groningen. Both pilot schemes extended early education programmes to cover all disadvantaged children in those regions by 2009/11. Because of this pilot, a majority of preschools in South Limburg introduced or intensified child-centred ECE approaches, in particular the method *Speelplezier*, while preschools in East Groningen introduce mostly a teacher-based ECE programme.

The target group has not yet been fully reached by ECE programmes. The original goal was to reach 50% of the target group within the period 2001/05 (OCW, 2001); the goal has then been raised to 70% for the period 2006/10 (OCW, 2006) – the pilot areas reached close to 100% by 2010. In a national survey among municipalities, Jepma et al. (2007) have found that at the ages of 2.5 to 4 years about 49-59% of the target group have been covered by ECE programmes; at 5-6 years the coverage has been about 54-85%. Given the earlier focus on urban areas, the coverage is much higher in bigger cities than in smaller cities and on the countryside.

A child who turns four can attend the first grade - nearly all children attend school in that grade, as of the age of 5 schooling is compulsory; often grade 1 and 2 form a combined group of pupils. During this initial stage of elementary schooling, children start to learn in a playful and simple manner to count, write and solve easy problems, and they advance their language and social skills in interacting with other pupils. Only from grade 3 onwards, when pupils are 6-7 years of age, is subject-based teaching introduced. Since 2000, municipalities are required to provide ECE to the majority of children who are at risk of falling behind in their school programme, at least when they attend the initial stage of elementary schooling (Nap-Kolhoff & Schilt-Mol, 2008).

Many municipalities have also introduced ECE programmes at preschool kindergartens and some even in daycare centres so as to stimulate children as early as possible. Those institutions are then supposed to apply the same ECE methodology as the elementary schools that most of their children attend afterwards. In such coupled partnerships, preschools and elementary schools communicate about the developmental progress of children who proceed from preschools to elementary schools and discuss joint extracurricular activities. These

programmes hence are deemed to offer a continuous approach with professional ECE staff at both institutions. Such an integrated approach is regarded in the literature as the most promising intervention to reach sustainable development gains (V. E. Lee & Loeb, 1995).

The Netherlands devotes about 0.4% of its GDP to pre-primary educational services, which matches the lower midfield of OECD countries, but remains below the EU target of 1%; Scandinavian countries, at 1.5-2% of their GDP, devote the biggest shares to pre-primary education (OECD, 2006). In the past, including most of the studied period, national funds addressing socially disadvantaged children were predominantly transferred to municipalities. Until the school year 2006/07, such targeting focused on the migration and educational background of parents; since then targeting focuses only on educational backgrounds.

Municipalities assigned part of these funds to investments in ECE programmes at selected institutions which covered areas where comparatively more children were at the risk of delayed Dutch language development as their parents spoke little or no Dutch with them. Since 2006/07, elementary schools have received their funds for ECE investments directly from the national government. Assignment of those funds to elementary schools depends on the pooled risk of enrolled children to endure educational delays because of any social disadvantages. For preschool kindergartens (and daycare centres) the amount of funding depends on how many groups and half-days the municipality deems necessary when placing special attention on disadvantaged children and how many it can afford to subsidize.

ECE funds are used to subsidize parental contribution fees, to equip centres with ECE education materials as well as to train staff in applying ECE didactic techniques. Fees are usually lower (or even non-existent) for children with disadvantaged backgrounds. Elementary schools invest their ECE funds to employ additional staff and thus reduce the child-staff ratios, to buy ECE education materials and to pay for ECE trainings for their staff. There are no tuition fees for elementary schooling in the Netherlands, except for private schools. Only preschools using an officially recognised ECE methodology can receive public funding.

A number of didactic approaches have been certified by the Dutch government as effective ECE programmes. Conditions for accreditation have been established that are based on what is in theory stimulating child development at preschool age (Inspectie van het Onderwijs, 2006; Nap-Kolhoff & Schilt-Mol, 2008; OCW, 2000, 2006).

The conditions are:

- A satisfactory intensity: covering the age group of 2.5 to 6 year-olds for ≥ 3 half days per week;¹⁶
- Appropriate and effective content: theory-based and well-structured didactic-pedagogical approach with evident positive effects on various development domains;
- Cooperation of preschools (and daycare centres) with elementary schools: continuation of the same approach throughout different ages;
- Child development is observed and registered;
- Enough qualified staff members per group: preschools and daycare centres with ECE programmes should not have more than 15 children per two professional staff member;
- Involvement of parents.

There are additional factors contributing to the success of ECE programmes but that are not all included in the official conditions; these are, for instance, measures to attract parents of the target population to send their children to a preschool that uses an ECE programme. It is unfortunate that most of the before-mentioned conditions are not very transparent in terms of measurable standards. Conditions are not fully met in all cases because of constrained resources, for instance, expenditures permit only three subsidized half-day sessions per week instead of four. Often the administrative capacity to control appropriately for the proper application of early childhood education at all locations is limited.

First assessments of *Speelplezier* by Bolt & Schonewille (2006) indicated that it has the potential to fulfil most conditions, though suggesting improvements with respect to the involvement of parents. In January 2009, the National Youth Institute (NJI) has assessed *Speelplezier* with respect to fulfilling the necessary conditions and granted an official ECE accreditation to the programme (Pennings, 2009).

Dekker, De Fijter & Veen (2000) provided a policy guide which gave municipalities an overview of the conditions as regulated in 2000. A recent update of this policy guide sub-classifies integral ECE programmes, into programme-directed and development-oriented approaches, and keeps an inventory of which early childhood education and parent-child programmes received certification (Nap-Kolhoff, Van Schilt-Mol, & Vallen, 2008). However, both inventories indicate that

¹⁶ Municipalities reported that they provide an average of 2.8 half days per week (Jepma et al., 2007; Middleton, 2009). Four or more half days per week is regarded as the didactically necessary intensity to reach significant gains in development (Van der Vegt et al., 2007).

the empirical evidence on effects of the ECE programmes has been rather incomplete, besides assessments of compliance with the conditions of accredited programmes. The classification of Dutch ECE programmes into programme-directed and development-oriented approaches goes back to Leseman, Otter, Blok & Deckers (1998). They argue that all programmes are in part development-centred as they all use some form of observation and assessment follow-up system to better address the needs of the individual child.

According to Nap-Kolhoff et al. (2008), the most commonly used ECE programmes in the Netherlands are *Piramide* (50% of ECE preschools) and *Ko-Taal* (15%), which are examples of programme-directed approaches; *Startblokken van Basisontwikkeling* (14%) and *Kaleidoscoop* (7%) are examples of development-centred approaches. Despite the fact that these programmes differ substantially, for instance, in the number of development domains that are addressed, the involvement of parents, the quality assurance and the scientifically expected effectiveness, programmes taken as a whole are somewhat similar once they qualify for official accreditation (Nap-Kolhoff, Schilt-Mol et al., 2008).

Speelplezier resembles to some extent the ECE approach *Kaleidoscoop*, which is based on the American *High-Scope* programme. It has a strong focus on the individual child and play activities and uses an observational record to keep track of the child rather than using a set of development tests. Children are clustered in smaller subgroups and receive special attention from the teaching staff according to their level of development each day in childcare. Teaching focuses on personifying, intervening and inspiring activities. Similar to the structured, teacher-based *Piramide* approach is the fact that it uses topic folders to guide the daily activities done with the children (Pennings, 2009). *Speelplezier* also puts a focal point on language development in its activities.

Institutions that do not use a distinctive professional ECE programme (yet) are typically referred to as 'regular' institutions. Henceforth, I refer to a *child-centred ECE preschool* as a preschool that has arranged at any time until the school year 2007/08 training for (part of) its staff to implement, for example, *Speelplezier* as child-centred ECE method.

Despite weak evidence on effects of Dutch early childhood care and education provisions in general, my hypothesis is that child-centred ECE institutions that are reaching out to the target group of disadvantaged children lead to significantly better outcomes for children than any alternative teacher-based ECE or regular preschool arrangements.

3. Literature discussion

There is still little evidence available about the effectiveness of early childhood care and education provisions, in particular in differentiating effects by the quality of such provisions. Quality evaluations of early childhood care and education investments can be classified into two types, those studying the processes and those studying structural quality aspects (see, e.g., Marshall, 2004). In this regard, effect studies of ECE programmes assess mainly the quality of the processes.

The international evaluation literature on early childhood education is dominated by a few U.S. American (quasi)experimental long-term studies, for example, of the child-centred High-Scope Perry Preschool, the curriculum- plus teacher-based Abecedarian programme and the teacher- plus parent-based Chicago Child-Parent Centres (for a review see e.g. Barnett, 2008). Studies of these programmes provided convincing evidence that early childhood investments lead in fact to positive effects in the long run, such as a reduction in rates of grade retention, drop outs, crime and delinquency, and the length of stay in special education, as well as increases in achievement scores and high school graduation rates. These programme examples imply for the society a ratio of investment costs to benefits between 1:6 and 1:13 (Nores, Belfield, Barnett, & Schweinhart, 2005). A critical review of this evaluation literature, envisaging the compromises that occurred in the randomisation protocol, identified lower but still significant social rates of return – well above the historical rate of return on equity (Heckman et al., 2010).

However, despite providing convincing evidence of the positive effects of early childhood education, the studied ECE programmes are of rather small scales and study experimental, high-quality and strongly targeted programmes. Policymakers cannot rely fully on those studies when deciding about larger scale programmes for other contexts. Scaling programmes up, as is being done in the Netherlands, usually demands reducing quality somewhat to an affordable and feasible level, and it becomes more difficult to monitor and enforce sufficient quality in the ECE provisions.

In the United States, Head Start is an example of a much larger-scale ECE programme; it is sponsored by the federal government. Its effect evidence has been ambiguous. Earlier studies showed that Head Start improves the test scores of children at the age of 5 (Magnuson, Lahaie, & Waldfogel, 2006) while positive effects fade away over time, in particular for black disadvantaged children (Barnett, 1992; Currie & Thomas, 1995; V. E. Lee & Loeb, 1995). More recent studies, however, identify also longer term gains of Head Start. Garces, Thomas & Currie (2002) identify long-term gains for Head Start participants by their early twenties, in high-school completion, college attendance and earnings outcomes, while for African-American participants delinquency rates are lowered. In a review of the international literature, Ruhm & Waldfogel (2011) confirm positive

gains from ECE expansions at school entry, adolescence and adulthood, in particular for disadvantaged children.

Studying a summary index of young adult outcomes, Deming (2009) finds that the disappearance of long-term effects may be caused by the fact that test scores are not covering the full scale of skills. He identifies that socio-emotional development gains can be found, in particular at later outcomes. A growing body of evaluation research on the state-sponsored universal pre-K programmes confirms short-term gains of larger-scale ECE programmes, in particular for disadvantaged groups (see, for example, Fitzpatrick, 2008; Gormley & Gayer, 2005; Huang, Invernizzi, & Drake, 2012). Esping-Andersen et al. (2011) argue that early development gains that may result from the experience of higher childcare quality such as early childhood education, as found for example for disadvantaged children in the U.S., may not last because the higher quality of attended childcare may subsequently not be matched by sufficient quality of schools to sustain those gains.

Evidence on the particular type of ECE programmes remains limited. Studies have shown that early non-parental language interventions at preschools, as in the case of the studied Dutch ECE programmes, can have a positive effect on the language skills of ethnic minority children, as discussed, for example, by Monte, Xian & Schweihart (2006). Their international comparison of preschool experiences of 10 countries shows that child-initiated, small group activities, which are consistent with developmentally appropriate practices, as applied in the *Speelplezier* programme, can promote active learning and school performance by the age of seven. This is in line with earlier studies, which compare the effectiveness of alternative ECE approaches, e.g. by Marcon (1992) and Stipek et al. (1998; 1995). They find that particularly child-centred approaches benefit children, as compared to mixed or teacher-directed approaches.

Barnett (2011) argues that large gains in other domains than cognitive development would likely require a balance of teacher-directed and child-initiated activities, including elements of dramatic play. Camilli et al. (2010) find in a meta-analysis of 123 studies of early childhood interventions that direct teacher instruction correlates strongly with cognitive child outcomes, more than child-oriented instruction. However, they stress that in the past many initial ECE interventions have been more strongly characterised by direct teaching, while child-oriented practices have been less advanced.

Haskins (1985) stresses that temporary intellectual gains from cognitively oriented ECE provisions could come at the cost of social behaviour deficits. Schweinhart et al. (1986) indicate that child-centred approaches can have longer lasting effects than those with direct instruction and show substantially better effects on non-cognitive outcomes, positive effects that would even increase over time, in particular in the socio-emotional, non-cognitive domain. However, these results

are based only on a limited sample. Huffman & Speer (2000) point out in this regard that more appropriate classrooms are related to better language and cognitive outcomes at the 3rd grade for Hispanic and Afro-American children when they attend a preschool that participates in the *Head Start - Public School Transition Project*. However, they cannot identify whether these effects are linked to the appropriateness of the classrooms or to the absence of overly regimented or highly disordered classrooms; they observed classrooms in which stronger emphasis has been put on basic skills and highly structured, direct teaching approaches.

Findings by Peisner-Feinberg et al. (2001) as well as Burchinal et al. (2010) suggest that high-quality classrooms and instruction are necessary to improve social and academic outcomes such as language, math and reading skills in pre-kindergarten ECE programmes, in particular for disadvantaged children such as those from low-income households. Heckman (2008) argues that, in particular, early childhood intervention programmes that put their attention on stimulating character development and motivation rather than focusing exclusively on the stimulation of cognitive development appear to be most effective.

Despite the limitation of the literature on the quality of ECE programmes, some studies attempt to shed some more light on the processes. Applications of environmental rating scales to assess a variety of such process quality factors by using composite indicators have become popular.¹⁷ The Dutch catalogues of conditions for ECE accreditation resemble simpler proxies of such quality assessments.

What evidence is available about Dutch ECE programmes? A review of the Dutch evaluation literature on ECE programmes by Nap-Kolhoff et al. (2008) has shown that effect studies on the major, maturing programmes are becoming outdated while studies on the newer programmes such as *Speelplezier* are still incomplete.

The most relevant evaluation study of ECE programmes¹⁸ has been done by Veen, Roeleveld and Leseman (2000), assessing the effectiveness of *Kaleidoscoop* and *Piramide*, two intensive ECE interventions (4 half-days per week), between 1995

¹⁷ See, for example, the environmental rating scales of the UNC / FPG Child Development Institute (2005): *ITERS* for infants and toddlers, *ECERS* for early childhood (e.g. Mashburn et al., 2008), as well as *FCCERS* for family daycare, and *SACERS* for children of school age. Applications of such scales can be found in the British EPPE study or the U.S. American NICHD study. Duncan (2003) provides an example of using the Observational Record of the Care-giving Environment *ORCE* as recorded in the NICHD Study of Early Child Care.

¹⁸ For an extensive discussion of characteristics of the programmes and their evaluation literature see Nap-Kolhoff, Schilt-Mol et al. (2008), as well as the meta-studies on effects of ECE attendance by Leseman et al. (1998) and the inventory of Van der Vegt et al. (2007).

and 1997, using a quasi-experimental research setup with two control groups and two experimental groups, one for each programme. Each group had a sample size of about 100 children; however, attrition was high, reducing the value of the statistical results. Simple correlations, controlling for earlier development outcomes and various background factors, show that children who attended an institution which applies one of the two ECE programmes show no significant difference in language and cognitive test scores as compared to the control groups. Litjens (2011) studies the impact of the 2000 policy impulse to extend ECE policies, matching children who entered elementary schooling before and after. She finds no significant effects of early childhood education, but some indication that ECE could lead to cognitive gains for native Dutch children with less educated parents.

Various reasons are put forward in the literature about what may have caused the lack of evidence on the effects of Dutch ECE programmes and that might confirm their theoretical support. Leseman (2002) argues that the intensity of early childhood care and education experiences has been insufficient; it is often not reaching the recommended four half-days per week. Driessen (2004) suggests that the professionalization of early education services has still not progressed sufficiently to reach significant improvements.¹⁹ Nap-Kolhoff et al. (2008) argue that the quality of services is still insufficiently documented so as to distinguish beneficial high-quality investments from non-beneficial low-quality investments. Jungbluth, Nap-Kolhoff & Rodigas (2011) point out that school test outcomes that are usually used for such evaluation research might be not good enough to differentiate effects on such a small scale.

To sum up, the evaluation literature on early childhood education programmes is still limited, e.g. by focusing on very specialised cases, and is dominated by U.S. American studies. The increasing use of ECE programmes, including in the Netherlands, is not backed up by sufficient research on their effectiveness. But neither does the scarce literature disprove their theoretical effectiveness.

4. Data description

To evaluate the impact of child-centred ECE programmes, I use a unique dataset (henceforth referred to as *Moelejaan* data) collected from 19 municipalities in the southern part of the Dutch province of Limburg, covering the first three grades of all publicly financed elementary schools in 2008/09 (Jungbluth, Rodigas, &

¹⁹ Professionalization refers to continuous integration between different levels of early childhood education and schooling (Driessen, 2004; Driessen & Doesborgh, 2003).

Bauchmüller, 2009).²⁰ This data thus covers three consecutive child cohorts born between the years 2001 and 2004, whose preschooling period was 2003-2006. This aggregates to a total of 16,679 child observations.²¹ These cover almost the entire child population aged 4-6 years in South Limburg in 2008/09.²²

The *Moelejaan* data provides rich information about these children from school registries and test administrations as well as a survey that has been conducted among parents. School registry information is available for (nearly) the whole population of children; next to birthdates and gender of every child it includes identification information, including home address, parental education levels and country of origin. Furthermore, in the autumn of 2008 all children took a questionnaire home to their parents; 11,088 children (~66%) returned at least a partially completed questionnaire to their teacher. The survey collected further information on family backgrounds as well as information on the child's history of early childhood care and education.

For the studied *Moelejaan* data, Jungbluth, Nap-Kolhoff & Rodigas (2011) show that preschool groups in which an ECE programme such as *Speelplezier* is applied tend to be have smaller groups, bigger rooms and more professional staff than regular preschools. On the other hand, staff in ECE preschools tend to be less experienced, which may be due to the fact that they are usually younger and have been recently recruited for ECE training.

I study short-term effects of child-centred ECE on school readiness at the beginning of elementary schooling. Most children who attend a preschool kindergarten with child-centred ECE subsequently attend a partnered elementary school applying such an approach, too. Even though ECE programmes are continued during the initial stage of elementary school, the lack of later outcome

²⁰ 207 of 210 elementary school locations have provided at least parts of the requested information (for a description of the fieldwork, please refer to Rodigas, 2009). Children in private schools (less than 0.3%), special schools (about 1.5%; CBS Statline, 2010) and children who attend primary schools across the border in Belgium or Germany are not included in this study. Estimates based on newspaper articles (e.g. Mariën, 2007) indicate that in 2006 about 5-10% of Dutch children living in the 19 municipalities of South Limburg attended an elementary school abroad. Belgian and German pupils who came to Limburg for the elementary schooling have been included in the data collection, but are not part of this analysis as they spent their preschool period abroad.

²¹ Observations where parents indicated that they wanted their children to be excluded from research have been dropped (~0.5% of all registered children).

²² The official number of children living in South Limburg in that age group was 16,616 according to the public administrations in 2009 (CBS Statline, 2009). No data is available on the small fraction of children participating in private and other specialised care.

measures in the studied *Moelejaan* dataset restrains me to study only the impact of the application of ECE programmes in the preschool kindergartens.²³ If I find short-term effects of preschool investments, they would prospectively be stronger when studying later child outcomes. Duncan et al. (2007) show in a meta-analysis that such earlier cognitive measures of school readiness can be strong predictors for later school performance.

To assess children's cognitive and language development, Dutch schools use nationally standardized tests of arithmetic understanding and Dutch language skills at half-yearly intervals. I use raw test scores of those cognitive and language tests taken in the middle of the second school year (January to February). Those M2 tests scores are nationally comparable, e.g. to studies based on PRIMA data from the Dutch national cohort study, and are available for more children than any other collected tests; data for both tests is available for 7,862 children.²⁴ To facilitate comparability of test scores with other studies, I have standardized scores for each test to means of 100 and standardized deviations of 15. Scores are normally distributed; see Figure III-1 for frequency distributions of the test scores and Table III-5 for a direct comparison of the original and standardized scores.

There was no central register or inventory of ECE applications in South Limburg at the time of the *Moelejaan* data collection. I enriched the *Moelejaan* data with ECE information on preschools from a survey among all 19 municipalities in South Limburg that I collected in early 2009. In this survey I investigated which preschools use an early education programme and when those programmes have been introduced. I cross-validated this information by doing an inventory of municipal documents and reports stating any information about preschool provisions and linked it to the preschool attendance details as stated in the initial survey among the parents whose children are covered in the *Moelejaan* data.

The studied child cohorts attended preschools by and large in the period 2003 to 2009. During that period six of the 19 municipalities in which data was collected had at least one preschool kindergarten applying *Speelplezier*; whereas nearly 90% of the *Speelplezier* applications occurred in two municipalities. In eight of the 19

²³ The gathering of test outcomes for all children at grade 3 and above and the inventory of ECE applications at the initial stage of elementary schooling are not yet completed; they are part of future waves of the *Moelejaan* data collection; a longitudinal set of observation of the studied children will then be available.

²⁴ The M2 tests are available for children who were in the second and third elementary school year at the time of the data collection, while children who attended the first school year were not tested yet at that time. Test data is available only for a maximum of two thirds of the children.

municipalities there were investments into other ECE programmes, however, at a much smaller scale.

Preschool kindergartens often apply ECE programmes only in part of their child groups. As I cannot identify which specific group a child attended, I measure average effects across ECE and *regular* groups of an entire kindergarten. I assume that there are positive spill-over effects from applying ECE in some groups to children in other regular groups as teachers in a location exchange their experiences. Moreover, I measure average ECE effects across all attended preschool days as I cannot identify for how many of the days of attending an ECE group a child was exposed to ECE programme elements. ECE programmes are usually not applied during all days of childcare. I assume that even if some of children's preschool days consist of *regular* childcare supervision, there are still spill-over effects from early education activities onto other days. Averaging across groups and days may imply that I underestimate the true effect of ECE programmes by diluting ECE effects when attributing them to children who did not benefit from it. Yet, the expansion of child-centred ECE may introduce variability that is attributable to other unobserved quality factors such as, for example, the quality of the local preschool director. Such added variability may help identify an effect that otherwise could be absent. Hence, the direction of any conceivable bias resulting from averaging across groups is not apparent upfront.

According to the survey among parents, about 78% of children attended a preschool kindergarten in South Limburg. They spent on average 153 half-day sessions in preschool kindergartens.²⁵ The survey provides, for 61% of all children, information on the length of preschool kindergarten attendance (for an overview of early childhood care and education attendance statistics for the studied sample, see Table III-7).

I reduce the studied child population to a sub-sample for which parents have responded to the survey and where I can link the provided information to a particular preschool kindergarten, for which I could find and cross-validate sufficient ECE information in my survey among municipalities. For about 44% of the children who attended a preschool, I could identify whether their preschool kindergarten used any or no ECE programme at the time when the child attended; hence ECE information is available for 27% of all studied children in South Limburg. Within this sub-sample for which information is available, 78% attended

²⁵ This is roughly in line with figures from the nationally representative cohort study PRIMA for 2004/05 that showed a preschool attendance rate of 75% with an average length of 131 half-days; the attendance rate for the unweighted PRIMA sub-sample for Limburg was 83% with an average of 146 half-days of preschooling (for more information about PRIMA, see Driessen et al., 2006).

a regular preschool without any ECE application, 11% attended a preschool kindergarten that runs a child-centred ECE programme (incl. *Speelplezier*), 7% attended a preschool with a teacher-based ECE programme, and another 3% attended any unspecified ECE programme.

5. Methodology

I want to measure the effect of child-centred ECE on child outcomes at the beginning of elementary schooling relative to any alternative preschool option. More precisely, I consider the effects on the child outcomes of language and cognitive development at age six after having participated at age two to four in some form of publicly provided preschool. I am not taking into account whether a child also experienced ECE at the beginning of the elementary school, up to the moment they took the test, but presume that they do, as preschools that apply an ECE programme are usually linked to elementary schools that also run such a programme.

I cannot exactly measure the children's development before or at the change from preschool to elementary school – early test outcomes are incomplete and less reliable child development indicators. The studied test outcomes thus reflect the full effect of early education at the preschool, including any indirect effect of preschooling that comes from the first months of a well-chosen, better quality elementary school trajectory that may be the result of improved preschool outcomes in the first place.

I test the impact of attending a preschool that used a child-centred ECE programme such as *Speelplezier* (*treatment group*) versus a group of children who attended either a preschool with alternative teacher-based ECE programmes or a regular preschool (*control group*). By looking only at children who attended any preschool of any type and by assuming it is likely that parents are little informed about the particular types of applied ECE programmes, I can disregard such unobserved selection biases that are due to parental choices for preschooling at the outset. This strategy helps in reducing potentially large biases on pre-treatment observables and thus in making stronger causal claims about the effects of child-centred ECE versus any other preschool investment.

I first compare the simple mean differences of the two assessed test outcomes; then I assess the problem of missing data bias and suggest the creation of a non-response weight to counteract it. Thereafter, I provide an Ordinary Least Squares (OLS) estimation of differences between the treatment and control groups while accounting for various background factors. In this OLS estimation the child i 's outcome (CO_i) is linearly determined by an average baseline score β_0 and varies according to the child's preschool history (I_i) and individual characteristics (X_i), as well as to family background factors accounting for part of the child's initial

development (F_i), plus a general error term, as shown in equation (III.1), X_i and F_i are assumed to be exogenous.²⁶ For an overview of descriptive statistics on the variables used, please refer to Table III-6.

$$CO_i = \beta_0 + \beta_1 X_i + \beta_2 F_i + \beta_3 I_i + \varepsilon_i \quad (III.1)$$

Despite relating the child outcome to the type of preschool care, I account for other care provisions. The collected data allows me to also cover part of the variation in initial investments at the period prior to children's preschool attendance. I therefore also include an indicator for the length of any centre-based daycare experience (that usually starts much earlier than preschooling) and for the estimated time a child was cared for at home by parents, other family members or a nanny. The duration of all three childcare arrangements is measured by the total number of half-day sessions.²⁷

As child characteristics, I control for gender, the child's age when taking the test in the middle of the second grade and for belonging to an ethnic minority in terms of having at least one parent who was born abroad. The family background factors account for single-parent households, the number of children per household, mothers' employment status at the time of the survey (the children's initial stage at elementary schools), the educational level of the parent who was educated the most, as well as income per household. Additionally, I exploit information collected in the survey about the cultural capital in the household with respect to language development. Therefore I use information on the father's Dutch language skills and the quantity of books for children available in the household.²⁸

Errors are clustered at the elementary school group level to account for potential nesting of types of children per group or systematic variation in the way children have been tested (=365 clusters); as a result I get more consistent standard errors. Potential biases caused by non-responses and missing information are accounted for by using weights for each observation in the regression.

5.1. Accounting for missing treatment information

²⁶ The exogeneity assumption implies that there is no unobserved ability, which is intergenerationally correlated. Any such correlation, e.g. with respect to parental affinity towards investing in early childhood that could reflect also parental skills to address the child's needs, would lead to biased OLS results.

²⁷ A typical preschool kindergarten session in the Netherlands lasts about 2-3 hours for any visited half-day (see OECD Review Team, 1999).

²⁸ Dutch language skills of mothers and fathers are highly correlated. I therefore include only the indicator for fathers' language skills which varies more than mothers' language skills.

Information about ECE attendance is not available for all children, as not all parents responded to the survey that I use to identify preschool attendance and not all individual ECE histories could be reconstructed using additional survey information on municipal preschool provisions. I would need to assume that this information is missing at random to prevent any bias in my results. Many cross-sectional studies using survey data tend to neglect this issue of item non-responses, as they either assume that it occurred randomly or because the issue of missing data occurs only in some explanatory factors.

But if, for instance, those who do not respond to the survey are overrepresented among those who did not attend any ECE preschools, my results could turn out to be underestimating the true returns to ECE attendance (see e.g. Rässler & Riphahn, 2006). I cannot neglect such potential non-response and missing variable bias as there is higher non-response for disadvantaged children (see Table III-9 in the appendix). ECE programmes are targeted at disadvantaged children and are meant to benefit them above average. Hence, I may underestimate the true effect of ECE preschools if I do not control for this lower representation of disadvantaged children in the studied sub-sample.²⁹

There are systematic differences between the full and the studied sub-sample in a number of descriptive statistics for the control variables on the individual child and family background as used in the linear regression. Univariate t-tests for the sample averages (see Table III-9) show that the studied sub-sample has a significantly lower share of boys, older children and ethnic minorities; of children with single parents and of unemployed mothers; of children with lower educated parents and with fewer books at home; and of those whose fathers have lower Dutch language skills and who live in poorer households. A multivariate probit regression (see Table III-10) of the sub-samples over the same child and family background factors confirms significant differences for children of ethnic minority and with unemployed mothers, mid-level educated parents and different numbers of books at home, but does not confirm the other differences found in the univariate tests. Yet, children of advantaged backgrounds that show significant differences allegedly benefit less than the average from child-centred ECE experiences. I may underestimate the true treatment effects if disadvantaged children are underrepresented in the studied sub-sample (see e.g. Rässler & Riphahn, 2006).

To account for any potential bias that is due to item nonresponse, I create a sampling weight that outweighs the bias which could result from using the sub-

²⁹ As I estimate the effects of child-centred ECE over any alternative preschool option, including teacher-based ECE preschools, this under-estimation may only be partial.

sample. The *Moelejaan* dataset has rich information available for a large share of the population. Weights are created in two steps, following the procedure as suggested by Little (1986).³⁰ First I generate propensity scores with the help of a logit model, the dependent variable being a response dummy to differentiate the two samples. Figure III-2.a and b (in the appendix) stress that the calculated propensity scores have a large common support and accordingly provide a good balance between the full sample and the studied sub-sample. In the second step, I use inverse values of the calculated propensity scores as weights for individual respondents.³¹

5.2. Group differences

Table III-1 shows summary statistics of the standardized test outcomes across two groups of children: 1) children who attended a regular preschool without any ECE or with a teacher-based ECE application, and 2) children who experienced a child-centred ECE preschool, e.g. one applying the *Speelplezier* approach.

Table III-1: Summary statistics on child outcomes

	Obs.	Mean	Std. Dev.	Min	Max
Dutch language test at middle of 2 nd grade					
Alternative preschool attended (teacher-based or no ECE method)	1199	102.63	12.99	37.82	123.51
Preschool with child-centred ECE attended	122	102.21	13.32	64.75	123.51
Cognitive test at middle of 2 nd grade					
Alternative preschool attended (teacher-based or no ECE method)	1199	102.61	13.29	43.92	127.06
Preschool with child-centred ECE attended	122	100.01	17.04	37.99	124.09

Note: re-weighted to account for non-responses or missing data in studied sub-sample, see section 5.1.

In particular for the cognitive test, child-centred ECE experiences are associated with somewhat lower average child outcomes. Only for this test is the difference marginally significant (t-test, $p=0.07$). The fact that child-centred ECE is associated

³⁰ For related discussions on the application of propensity scores as sampling weights in case of potential nonresponse biases, see also Yansaneh (2003), Wun et al. (2004), Harrod & Lesser (2006) as well as Potter et al. (2006).

³¹ To prevent that respondents with very low propensity scores receive excessively large weights that could inflate the variance of the survey, I exclude outliers, using a 0.01 cut-off level; 1 out of 1322 observations is dropped.

with lower average test outcomes could be a result of its targeting to disadvantaged children who tend to lower the average score. That the difference is larger for cognitive test scores could be due to the fact that child-centred ECE such as the *Speelplezier* approach explicitly focus on stimulating linguistic development rather than cognitive development.

5.3. Linear Regression

I first study the effects of child-centred ECE with an OLS multiple regression model. I assume that the relationships between explanatory factors and outcomes as well as a number of control factors are linear, which is in line with the majority of papers evaluating educational investments with cross-sectional data. For the treatment period, I include a simple and a quadratic term of the length of the preschool attendance to adjust for potential non-linearity in the relationship between preschool duration and child outcomes (see, for example, Landvoigt et al., 2007). And I control for various child- and family-related characteristics that are also deemed important for determining child outcomes.

Interaction effects between the treatment variable and the family background characteristics can indicate whether the programme is more beneficial to children who have disadvantaged backgrounds. I am interested in whether effects are heterogeneous, e.g. by gender, parental educational attainment, by single-parenthood or the child's origin.

I provide eight different OLS estimation models: a basic model per test child outcome (models 1 and 4) that includes only the attendance spells of different early childhood care experiences as well as other child characteristics and family background factors; the basic models extended by the treatment dummy for attending a child-centred ECE preschool (models 2 and 5); and the extended basic models including the treatment dummy as well as its significant interactions with any subgroup dummies (models 3 and 6).

To make any causal claims about the tested relationship I need to assume an exogenous treatment allocation process that is unrelated to the child outcome. If I_i is exogenously determined, i.e. the choice of preschool institution is at random, the OLS estimates are consistent and efficient (see Duncan, Magnuson, & Ludwig, 2004). OLS may overstate the effect if high-quality children are selected into high-quality preschools (Angrist & Krueger, 2001).

The studied ECE approaches have been mainly introduced as a result of the year 2000 policy to target disadvantaged children in larger cities. The new ECE regulation (*regeling VVE*) of the Ministry of Education, Culture and Science (OCW, 2000) gave a financial impulse to extend ECE application to some preschools, in particular in bigger municipalities to help better integrate its larger populations of

ethnic minority children into elementary schooling. Notwithstanding a general tendency to improve the quality and to extend the availability of preschool places, the studied period has seen no major policy changes affecting preschool kindergartens.

In the Netherlands many parents choose a childcare arrangement according to availability and to some degree according to the price. Often they are less aware of whether any preschool applies a child-centred ECE approach or not. There is some evidence from fieldwork and survey data that there is little awareness of the existence of particular ECE approaches such as *Speelplezier*. Of those parents who mentioned in the survey that their child attended a *Speelplezier* preschool and whose children can be matched to the truly applied ECE approach at the preschool, only 33% have actually experienced *Speelplezier*. Quoting of other ECE methods has a similarly low appropriateness, and there is no significant variation of those levels across background factors such as, for example, parental education level (see Table III-8 in the appendix).

The presence of such strong measurement errors in ECE variables from surveys might induce attenuation bias if not accounted for (see, e.g. Frost & Thompson, 2000), which might be a reason why some comparable studies have not found ECE effects yet. By using only a sub-sample of children for which I match children exactly to their preschool and its ECE history, I prevent dilution of my OLS regression.

Awareness of ECE applications may be somewhat higher among parents of a higher socio-economic background, but this difference can be accounted for by including parental factors in the analysis. Parents differentiate little between the quality of a childcare arrangement according to whether an ECE method is used or not. Those who send their child to a preschool kindergarten that uses an ECE method do not value the preschool quality higher than parents whose children attend a regular preschool.³²

Given that there is some exogeneity in the allocation of children into preschools applying different ECE approaches, I argue that parents do not care much about which ECE approach their children experience. Nonetheless, I will apply alternative estimation techniques, i.e. Propensity Score Matching and Instrumental Variables application, in order to deal with possible endogeneity.

³² The international literature confirms that preschool quality seems of lower importance in parental childcare choices than, for example, the price (see e.g. Blau, 1991; Blau & Hagy, 1998).

5.4. Propensity Score Matching (PSM)

As the assumptions about a linear parametric model specification and the random character of the allocation process might be too strong to withhold against any doubt, I also use two alternative strategies: Propensity Score Matching (PSM) and an instrumental variable approach.

The application of the propensity score matching technique is another way to do a counterfactual reasoning and thus to determine the average outcome the child-centred ECE attendees would have experienced had they attended an alternative preschool, which gives us the Average Treatment Effects for the Treated (ATT) (see Caliendo & Kopeinig, 2005, for a step-wise description of applying PSM). The advantage of PSM over a linear regression is that parametric assumptions of linearity can be relaxed; it allows for more robust predictions within the area of common support as it does not make any extrapolations outside of that area (see e.g. Cochran, 1963). Moreover, PSM focuses the researcher's attention on the direct comparability of treatment and control groups.

Propensity scores are balancing scores based on a number of observed characteristics from the overall population whose conditional distributions are independent of the assignment into treatment, as described by Rubin & Rosenbaum (1984) as well as Imbens & Rubin (2008). I reduce the whole *Moelejaan* data sample to a sub-sample of observations for which sufficient information about individual ECE experiences is available. I then match people who are on average alike in terms of their observable pre-treatment characteristics, including their attendance of a preschool kindergarten of any kind.³³ The only difference that is allowed is that the *treatment group* attended a child-centred ECE preschool while the *control group* did not, i.e. the latter experienced any other ECE programme or regular preschooling.

To generate the propensity scores, I use the same child and family background factors that I use as controls in the linear regression model. I have cross-sectional data and thus no baseline information about the initial development of the children studied and their parental background. I hence assume that the indicators, which I use for my propensity scoring, are stable over the preschooling period; only the treatment is supposed to differ.

PSM applications make the strong assumption that selection into treatment is based solely on observable characteristics, implying that all variables that significantly influence the assignment of child-centred ECE preschooling and the

³³ I use the Stata command `PSMATCH2` to implement full Mahalanobis matching and different direct and closest neighbour PSM methods (see Leuven & Sianesi, 2003).

studied child outcomes simultaneously are observed (Conditional Independence Assumption or CIA). I have a rich dataset of individual and family characteristics at hand to predict selection as well as possible. As discussed for the case of the linear regression, I argue that the cases were treated to a large degree at random. While the programme has been targeted towards preschools in areas where a potentially “weaker” child population lives, the targeting has been far from perfect – so it is likely to find children with similar characteristics among those who have been treated and those who have not been treated.

Standard errors need to be corrected, as PSM brings along additional variation beyond the usual sampling variation (Heckman, Ichimura, Smith, & Todd, 1998). Bootstrapping accounts for this problem, which is valid as long as the sample size is sufficiently large. When applying this method, random draws from the given sample are taken, with a large number of repetitions so as to create a large number of randomly reordered datasets. In this regard, bootstrapping relies on the statistical distribution as it is in the data; it does not have to make any parametric assumptions.

The area of common support assures that that children with the same propensity score have a positive probability of having both experienced a child-centred ECE preschool or any preschool alternative (Heckman et al., 1999). Treated observations which lie outside the common support are not matched and are excluded. While a match of control and treatment group is better if more information is used to calculate the propensity score, each additional variable exponentially increases the multivariate dimensions necessary to find a common support area.

Propensity scores ought to be based on a number of variables that credibly influence the participation in a child-centred ECE preschool as well as the test outcomes (Grilli & Rampichini, 2011; Smith & Todd, 2005). Extraneous variables that influence only participation could be included to increase the fit of the estimated model. However, an over-parameterised model can reduce the area of common support and thus reduce the number of successful matches; it can also increase the variance of the estimates.

Balancing tests, before and after matching, are used to check whether the propensity score optimally balances the treatment groups with the matched control group. Covariates are well balanced when the percentage bias after matching is less than 5% and the t-tests are non-significant (Grilli & Rampichini, 2011; W.-S. Lee, 2006). The bias is defined as the difference of the mean values of matching variables of the treatment group and the control group (before and after matching), divided by the square root of the average sample variance of the original full sample.

A too small common support can give some indication of selection bias due to excluded factors which are non-overlapping (Bryson, Dorsett, & Purdon, 2002). If the overlap is too small, then PSM may not be the most accurate method to apply; the external validity of estimation results from the overlapping sub-sample might be limited. Next to matching to the nearest as well as five nearest neighbours (1-to-1 and 1-to-5 matching), I also apply Kernel matching, which potentially produces a bigger common support area. For the Kernel procedure I make use of the whole common support and matches weighted by the distance between the propensity scores of the matched treatment and control group units (Bryson et al., 2002).

Despite correcting for biases caused by observables, propensity score matching cannot overcome potentially significant biases based on unobservables; omitting important variables could eventually even increase the resulting bias in estimated treatment effects (Heckman, Ichimura, & Todd, 1997). The CIA is in principle an untestable assumption that we cannot fully prove or disprove.

If I_i is endogenous, both OLS and PSM estimates are inconsistent. Endogeneity may occur if the choice of preschool institution is not completely random, e.g. if parents influence the choice of preschool institution in an unobserved way. If high-ability children are more likely to be sorted into such child-centred preschool institutions while we cannot identify their initial ability, OLS or PSM estimations might overstate the effects of such preschools. In that case, estimation by OLS and PSM will produce an upward-biased estimate of β_3 and will therefore not identify true causal effects of the daycare institution on children's outcomes.

5.5. Instrumental variable (IV) application

If preschool quality is endogenous, I expect β_3 in the estimated equation (III.1) to be smaller when applying the IV method rather than when estimating by OLS or PSM. If parents cannot truly influence the choice of preschool institution, the estimated OLS (or PSM) and IV coefficients are both consistent, but estimation by OLS or PSM is more efficient.

As an attempt to overcome potential biases that are due to unobserved factors, I introduce IVs that are correlated with the provision of child-centred ECE at preschools, but are unrelated to the child outcomes. All applied instrumental variables use exogenous sources of information that presumably provide an element of randomness in the preschool allocation process (Angrist & Krueger, 2001).

I use the following four instruments in a Two-Stages Least Squares (2SLS) estimation: 1) municipality dummies that indicate whether a municipality has already invested into child-centred ECE, 2) dummies for the management boards of the preschool kindergartens to control for a variation in ECE policies, 3) the

distance between homes and the nearest child-centred ECE preschool, 4) the distance to the nearest child-centred ECE preschool over the distance to any nearest alternative.

The validity of those four instruments as effective exclusion restrictions cannot be tested directly (Angrist et al., 1996); they need to be confirmed by theory. I argue in the following why I perceive those instrumental variables to be valid in providing exogenous treatment information. The strength of those IVs is later on assessed in terms of first-stage t-statistics.

The first source of instrumental variable information is related to the *municipality in which preschool kindergartens are located*. Only a small share of the 19 municipalities made investments in *Speelplezier* or other child-centred ECE methods during the studied period. And as parents cannot easily register their child at a subsidized preschool kindergarten in another municipality, choices with respect to ECE availability at preschool kindergartens are thus limited to some parents. As instrumental variables I thus include a dummy for whether municipalities use *Speelplezier* or not. The instrument might be questioned on the argument that parents who care about their children's education tend to elect politicians who adopt similar policies. However, given that ECE programmes are targeted to disadvantaged children, corresponding parents would care less about electing this type of politicians.

A similar source of instrumental information relates to the *management board of preschool kindergartens*. Management boards tend to promote a common approach across preschools towards using ECE. Some have invested strongly into ECE, e.g. *Speelplezier*, while others have not made any ECE investments at all yet - 3 of 10 boards have invested in child-centred ECE, covering preschools with about half of the children studied. Many preschools are independent from a higher level management board, and some management boards cover preschools in different municipalities. Parents seem to have little awareness not only of the variation in ECE applications but also of the ECE policy variation across management boards. Hence they have little preference for sending their child to a preschool under a specific management board versus another.

The third and fourth sources of exogenous information on treatment allocation are based on the geographic availability of child-centred ECE applications. According to the survey among parents, 78.3% favour a facility at a short distance from home. The *Moelejaan* data and the inventory of the preschool kindergartens include information about the postcodes of the child's household and the preschool's location. After having matched those postcodes to GPS coordinates, I calculated

distances between individual households and preschool locations.³⁴ I create the following two instrumental variables: the *distance to the nearest preschool kindergarten which uses child-centred ECE* as well as the *ratio of this distance over the distances to the five closest alternative preschool locations*. Both geographic sources of exogeneity build on the assumption that parents do not move to a housing location to be closer to their preferred preschool location.

However, it is not clear how parents would choose - the preference to access a preschool with presumably higher quality that is due to an ECE application is likely to be outweighed by an aversion of the weaker peer group in that area. This is because ECE investments are targeted towards areas with more disadvantaged children - ethnic minorities and children of lower educated parents.

As a result of such targeting, parents may consider the ECE applications not only as a sign of higher childcare quality but also as a label that indicates a more problematic peer group. Hence, targeting may provide signals that can lead to segregation. Due to such stigmatization of ECE preschools parents may be inclined to choose a regular preschool even if the quality of the provided childcare at such a preschool would be lower. On the other hand, if disadvantaged parents are on average less informed and knowledgeable of the existence of ECE applications, they may not stigmatize ECE preschools by its weaker peer group, or they may not have the resources to send their child to another preschool.

Moelejaan data indicates that preschools with larger distances from home are much less often chosen. The data also shows that less educated parents are more likely to choose a destination that uses an early education programme than parents with average or high education levels. Among all parents, the presence of an early education programme makes them less likely to choose a particular preschool.

The *Moelejaan* survey reveals reasons why parents chose a specific preschool investment for their child. It shows that the peer group is a less important argument when choosing a preschool than the quality of the childcare staff and the institution as well as the study materials being used. A cross-check across different socio-economic groups of parents shows that stigmatization happens to some degree; parents with more advantaged backgrounds tend to choose more in terms of peers at an institution. Still, across all groups the importance of ECE quality ranks higher than the social composition at an institution.

³⁴ Dutch Postcodes include a 4-digit number, referring to small sub-areas of municipalities, and a 2-letter combination that usually varies within the same street. Calculations for 2009 show, there are on average 17 private households per 6-position postcode. Postcodes have been matched to GPS coordinates that were tracked from free online maps. Distances are based on a simple Pythagorean formula and recalibrated into kilometre units.

A look at the social compositions of preschools with and without child-centred ECE applications shows that, on average, children at child-centred ECE locations come from households with lower incomes, fewer books and with fathers that have weaker Dutch language skills (see Table III-11 in the appendix). Yet, the differences for the two targeting factors – ethnic minority and lower parental education – are not or are only marginally significant.

6. Estimation results

In the following I discuss the output of the OLS estimation (see Table III-2) for the explanatory factors on the child's characteristics and family background, for the variation according to different childcare provisions and, finally, the variation according to the child-centred ECE method and its subgroup interactions.

6.1. OLS estimates

Table III-2: OLS regression output

Dependent variable	Dutch language skills			Cognitive skills		
Estimated OLS Model	1	2	3	4	5	6
ECE treatment indicator						
Child-centred ECE (attendance dummy)		0.76 [1.42]	-0.87 [1.49]		-2.08 [2.04]	-3.69* [2.01]
Subgroup interaction effects						
Child-centred ECE X Highest par. education: level 1			13.43*** [3.90]			13.21*** [3.05]
Other childcare indicators						
Total length of preschool attendance (in units of 10 half-day sessions)	-0.22* [0.12]	-0.22* [0.12]	-0.20* [0.12]	-0.10 [0.10]	-0.09 [0.10]	-0.08 [0.10]
Total length of preschool attendance (squared term)	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]
Total length of childcare attendance (in units of 10 half-day sessions)	-0.03 [0.03]	-0.03 [0.03]	-0.02 [0.03]	0.00 [0.03]	0.00 [0.03]	0.00 [0.03]
Total length of childcare at home (in units of 10 half-day sessions)	-0.01 [0.01]	-0.01 [0.01]	-0.01 [0.01]	0.00 [0.01]	0.00 [0.01]	0.00 [0.01]
Child characteristics						
Gender (boy = 1)	-3.39*** [0.76]	-3.41*** [0.76]	-3.61*** [0.75]	-1.88*** [0.71]	-1.82** [0.71]	-2.02*** [0.70]
Age at time of testing (in years, middle of 2 nd grade)	5.24*** [1.03]	5.23*** [1.03]	5.16*** [1.02]	6.46*** [0.98]	6.50*** [0.98]	6.44*** [0.98]
Ethnic minority (min. one parent born abroad = 1)	0.35 [1.34]	0.39 [1.33]	0.59 [1.32]	0.27 [1.36]	0.14 [1.34]	0.34 [1.33]

Family background factors						
Parenthood (single parents = 1)	-0.28 [1.84]	-0.30 [1.84]	-0.49 [1.77]	-1.82 [1.98]	-1.76 [1.96]	-1.95 [1.98]
Number of children in household	-1.27** [0.50]	-1.26** [0.50]	-1.17** [0.50]	-1.05* [0.56]	-1.08* [0.55]	-1.00* [0.55]
Mother's unemployment	-1.16 [3.67]	-1.22 [3.68]	-1.33 [3.58]	-4.37 [2.75]	-4.21 [2.70]	-4.33* [2.52]
Highest parental education: level 2 (middle / professional educ.)	3.79*** [1.43]	3.79*** [1.42]	4.99*** [1.45]	2.54** [1.29]	2.56** [1.29]	3.74*** [1.29]
Highest parental education: level 3 (higher professional / acad. educ.)	4.82*** [1.48]	4.78*** [1.48]	6.03*** [1.50]	2.42* [1.36]	2.52* [1.37]	3.74*** [1.37]
Number of children's books at home (1 unit = 5 books)	1.16** [0.45]	1.16** [0.45]	1.20*** [0.45]	0.99** [0.48]	1.00** [0.48]	1.04** [0.47]
Father's Dutch language skills (5 point Likert scale)	1.69*** [0.60]	1.71*** [0.61]	1.70*** [0.60]	1.70*** [0.56]	1.64*** [0.57]	1.63*** [0.57]
Net monthly hh-income (logarithmic scale)	0.78 [1.13]	0.79 [1.12]	0.62 [1.12]	2.82** [1.20]	2.78** [1.19]	2.61** [1.18]
Constant	62.02*** [7.16]	61.96*** [7.14]	61.16*** [7.04]	49.93*** [7.03]	50.10*** [7.04]	49.31*** [7.02]
Observations	1321	1321	1321	1321	1321	1321
Adjusted R ²	0.08	0.08	0.09	0.07	0.08	0.08

Note: Cito Dutch and Cognitive tests are taken in the middle of the 2nd grade, standardized scores (mean 100, standard deviation 15) - see Figure III-1 and Table III-5; * significant at 10% level, ** at 5% level, *** at 1% level; clustered at individual elementary school group level (robust standard errors in parentheses); analysed sub-sample re-weighted to covariate distribution of whole child population in South Limburg.

Even though the estimated models use a common set of factors about child characteristics and family background, they can explain only about 8% of the variation in the data (see Adjusted R²). Despite having a rather big sample size, a number of explanatory factors show no significant correlation to the two assessed child outcomes. In particular, single-parenthood and belonging to an ethnic minority shows no significance for any of the estimated regression models. More children living in a household has adverse effects on child outcomes. Girls strongly outperform boys, as do older children.

The other explanatory factors are significantly correlated with the child outcomes and show the expected signs - for an extensive review on the link between child outcomes and various background factors see, for example, Björklund & Salvanes (2011). The estimated negative effects of the mother's unemployment and positive effects of household income are only significant for cognitive test scores, which could indicate that language development is insensitive to household investment – an interesting finding that needs further research attention.

Factors of the household's cultural capital show the expected signs. Parental education has a significant positive relationship with both development indicators; effects are particularly large for Dutch language skills. The quantity of children's books in the household (grouped in categories of 5 books), a proxy for literacy environment and learning culture in the household, has a positive and significant relationship with both test outcomes. The same is the case for fathers' Dutch language skills (which are strongly related to the mother's language skills in the first place).

When looking at the care arrangements that a child has experienced before elementary schooling, we find only a few significant associations. Attendance at preschool kindergartens has the most sizable effect, but the negative association is still negligibly small – substantial changes in the experienced preschool term would be needed to get any sizable effects. Childcare at home shows some negative association with Dutch language skills, and centre-based daycare shows a negative association with cognitive skills; but these effects are even smaller. Overall, the negative signs of those associations are unexpected.

Many estimates for the periods of care attendance are insignificant, and those that are significant are rather small. This may be because formal childcare arrangements at home or centres still constitute a rather small share of children's experiences, while home-based care may be underreported. Likewise, a lack of significant estimates could be due in part to bigger measurement errors in those indicators.

Turning to the crucial effect estimates of the ECE programmes, I find that child-centred ECE experiences show no significant main effects for both Dutch language and cognitive test scores, compared to children who attended any other preschool with teacher-based ECE or no ECE approach. However, when looking at the effects of child-centred ECE across a range of subgroups of children, I find that such an experience has a powerful positive effect on both child outcomes for those children who have less educated parents.³⁵ In the studied period, low parental education was one of the risk factors that ECE was targeted to compensate for, in addition to belonging to an ethnic minority. With the interaction term included, the main ECE effect turns negative for the cognitive outcomes, which might point towards child-centred ECE not being able to address the needs of children with better educated parents, e.g. due to insufficient staff qualifications.

I also check whether the estimated effects are affected by any outliers, so I run those models again as quantile regressions (excluding the clustering option).

³⁵ Only significant effect differences between those subgroups and the main group are presented; the whole set of effect estimations is available upon request.

Quantile regressions are estimates of central tendency where median values are resistant measures that are not as greatly affected by outliers as are the means. The sensitivity check shows similar levels for significant estimates as in the OLS regression. The zero mean effect does not hide any impact at some point of the distribution. Yet the quantile regression indicates that outliers might influence the effect estimate for the subgroup of children with less educated parents. The interaction effect for language outcome is somewhat lower (9.58) and thus a still remarkably high but more reasonable level; the main effect for the language outcome stays the same. For the cognitive outcome, the interaction effect does not change, whereas the main effect loses its significance (it is significantly negative in the OLS regression).

When rerunning to the OLS regression without applying weights as described in section 5.1, the main effect for child-centred ECE turns out to be small but significantly negative for the language outcome, whereas all other estimates do not change much. This backs up my assumption that estimation without controlling for item non-response in the treatment variable could lead to a downward bias in the estimated treatment effects.

6.2. PSM estimates

Are those results confirmed when applying propensity score matching? Table III-3 provides an overview of average treatment effects on the treated (ATT) for three matching techniques: 1) nearest neighbour matching, 2) 5-nearest neighbours matching, and 3) Kernel matching. The significances of the differences between the group of treated and the group of matched controls are tested by applying the bootstrapping method.

Table III-3: Propensity score matching

	Treated		Controls		Bootstrapping		
	Mean	N	Mean	N	Average treatment effect on treated (ATT)	Std. Error	P> z
Nearest Neighbour matching							
Dutch language test (M2)	102.500	122	103.062	1199	-0.562	2.167	0.80
Cognitive test (M2)	99.998	122	104.184	1199	-4.186	2.270	0.07
5-Nearest Neighbours Matching							
Dutch language test (M2)	102.500	122	103.379	1199	-0.879	1.606	0.58
Cognitive test (M2)	99.998	122	102.972	1199	-2.974	1.889	0.12
Kernel matching							
Dutch language test (M2)	102.500	122	102.160	1199	0.341	1.358	0.80
Cognitive test (M2)	99.998	122	102.325	1199	-2.327	1.565	0.14

Note: Cito Dutch language and cognitive tests are taken in the middle of the 2nd grade, with standardized scores (mean 100, standard deviation 15) - see Figure III-1 and Table III-5; average Treatment Effects for Treated (preschool attendees); treated and control group within common support; and bootstrapping on ATT with 100 replications.

All three matching methods achieve substantial percentage reductions in the absolute standardized bias of the vast majority of explanatory variables before and after matching; balancing tests show that after matching there are no longer significant differences between the treated and non-treated group in any variable. The matched sample is sufficiently large to have statistical power; the treatment group size is above 100.

Each of the three matching methods results in insignificant differences between the treatment and control groups, which is in line with Smith and Todd (2005), who stress that, in practice, the choice of matching method often turns out to make little difference. Still, the joint consideration of several approaches offers a way to check the robustness of the treatment estimates. The PSM results confirm the results from the linear OLS regression, which also show no significant main effects of child-centred ECE.³⁶

6.3. IV estimates

In the following section, I use different sources of exogenous information on the allocation process of child-centred ECE in an attempt to correct for potential unobserved biases. I present the results of two stage least squares estimations using the previously mentioned instrumental variables. First-stage results allow an assessment of the strength and validity of my chosen IVs in determining the allocation of child-centred ECE. A whole range of IV combinations is tested for best fit.

All four indicators are shown to be valid instruments when used separately.³⁷ However, a combination of all four leads to over-identification problems and the geographic indicators on distances between homes and preschools does not predict

³⁶ As the areas of common support for the different child outcomes are already small, in particular for the second testing, I cannot study subgroups in more detail with sufficient statistical power.

³⁷ A sensitivity test on whether the full set of IVs is jointly strong enough to predict each of the two potentially endogenous repressors – child-centred ECE attendance and its interaction with the group of children with less educated parents, shows that the joint set of IVs can significantly predict each of the two regressors (following the procedure described by Angrist & Pischke, 2008, p. 218).

significantly the treatment variable when combined with the other two IVs, which are the indicators of whether the municipality and the management board of a preschool invest in child-centred ECE. Therefore, I present only the output for the investment indicators (see Table III-4.a).

Second-stage results show the estimated coefficients in the OLS versus the IV models: see Table III-4.0. In addition, I run post-estimation tests: Wooldridge's test assesses the exogeneity of the instrumented treatment variable; while Hausman's consistency and efficiency test shows whether the IV estimates are preferred over the OLS estimates because of systematic difference (see Table III-4.c).

Table III-4: 2SLS regression output, applying all 4 instrumental variables

a) 1st stage results

Dependent variable	Dutch language test (M2)	Cognitive test (M2)
Estimated IV model	2SLS (2)	2SLS (5)
Instrumental variables used for child-centred ECE attendance dummy		
Municipality invests in child-centred ECE	0.347*** [12.67]	0.347*** [12.67]
Management board of preschool invests in child-centred ECE	0.019*** [2.72]	0.019*** [2.72]
1. Joint significance of the instrument(s) in the first stage; critical values $F > 10$	$F = 19.76$ / $p = 0.00$ IVs strong	$F = 19.76$ / $p = 0.00$ IVs strong
2. Over-identifying restrictions for 2 nd stage regression: Sargan (1958)	$\chi^2 = 0.67$ / $p = 0.41$ IVs valid, No over-identification	$\chi^2 = 1.01$ / $p = 0.31$ IVs valid, no over-identification

Note: Robust standard errors in parentheses; * significant at 10% level, ** at 5% level, and *** at 1% level. A critical F value for the test on joint significance of the IVs (see test 1) is 10 (as stated, for example, in Baum et al., 2003); the test for the over-identification is based on an estimation without robust standard errors and weights are forced. It checks the coherency of the chosen set of instruments rather than their validity (cf. Parente & Santos Silva, 2012).

b) 2nd stage results

Dependent variable	Dutch language test (M2)		Cognitive test (M2)	
	OLS (2)	2SLS (2)	OLS (5)	2SLS (5)
Estimated IV model				
Child-centred ECE (attendance dummy)	0.76 [1.42]	0.10 [2.98]	-2.08 [2.04]	-2.75 [3.36]
Other childcare indicators	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES
Family background factors	YES	YES	YES	YES
Constant	61.96*** [7.14]	62.01*** [7.11]	50.10*** [7.04]	50.16*** [6.98]
Observations	1321	1321	1321	1321

Note: Cito Dutch language and cognitive tests are taken in the middle of the 2nd grade, (standardized at mean 100 and standard deviation 15) - see Figure III-1 and Table III-5; Robust standard errors in parentheses, clustered at individual elementary school group level; * significant at 10% level, ** at 5% level, and *** at 1% level.

c) Post-estimation model tests

Dependent variable	Dutch language test (M2)		Cognitive test (M2)	
	2SLS (2)		2SLS (5)	
Estimated IV model	2SLS (2)		2SLS (5)	
3. Wooldridge's test of exogeneity of instrumented treatment variable(s)	F=0.07 / p=0.80 Exogenous		F=0.06 / p=0.80 Exogenous	
4. Hausman test for systematic differences between 2SLS and OLS estimation model	$\chi^2=0.01$ / p=1.00 no difference		$\chi^2=0.04$ / p=1.00 no difference	
Interpretation of IV results	Prefer OLS		Prefer OLS	

Note: I test for each of the two-stage least squares models whether it is better than an ordinary least squares specification, i.e. I use a Wooldridge test (1995) to see whether OLS are providing consistent estimates because the instrumented variables are actually exogenously determined; for this option I force weights. And I use a Hausman test (1978) to check for systematic differences in the consistency and efficiency in the specifications of each IV and OLS estimation model.

The 2SLS estimations reproduce the results of the OLS and PSM in terms of finding no significant effects of child-centred ECE over alternative preschool options. Whichever 2SLS specification I use, Wooldridge's tests of the exogeneity of the IVs show that I do not need to worry about endogeneity in the treatment variable. And Hausman's tests for systematic differences between the OLS and 2SLS show that none of the tested 2SLS models is systematically more efficient than an OLS model. However, the chosen IVs might be valid but not strong enough to produce systematically more efficient 2SLS estimates, i.e. due to too high standard errors for the 2SLS estimates.

Unfortunately, the four IVs do not have sufficient strength to predict a second treatment factor, i.e. interaction terms of the child-centred ECE with any subgroup.³⁸ In some cases two (or more) of the IVs show a significant relationship in the first-stage results (even though the F-statistics for the joint significance is below 10), and Wooldridge's endogeneity test confirms that the treatment variables are endogenous.

The positive direction of the effects of child-centred ECE on language development of children with less educated parents, as found in the OLS estimation, are significant at a 10% significance when using all four IVs jointly or when combining only the first IV (municipal investment in child-centred ECE) and the third IV (ratio of distance to nearest child-centred ECE over alternative preschools). However, for those models the size of effects is far beyond any reasonable level, which could be a result of the lack of strengths of the IVs to determine both treatments.

7. Conclusions

While there is mounting evidence that enriched early childhood investments can particularly stimulate disadvantaged children, little is known yet how such investments should look. There is a growing debate about what a good early childhood education programme should be: structured, curriculum- and teacher-based, or instead holistic and centred on the individual child. In this chapter, I investigate whether children who attended a preschool with a child-centred ECE programme show more development gains than those who attended alternative preschool options. And I study whether such a programme can level the playing field for children who are at risk of not being developmentally ready for elementary school when they enrol.

The estimates show that child-centred ECE programmes, such as the newly certified early childhood education programme *Speelplezier*, have not yet led to significant positive (nor, for that matter, negative) effects on language and cognitive development. These results are in line with the given literature that finds little evidence for the effectiveness of Dutch early childhood education and thus challenge again the positive effects on cognitive and school outcomes found in international studies on individualised ECE approaches, as reviewed, for example, by Camilli et al. (2010).

The evident lack of overall main effects of child-centred ECE experiences cannot be easily generalised to other settings. As a matter of fact, I have used a restricted sample of children to estimate the model. In particular, I have used data for the

³⁸ 2SLS Estimation output for those interaction models is available upon request.

specific case of South Limburg. This region faces fewer problems with integrating ethnic minority children into early childhood education institutions than the rest of the Netherlands. It rather observes disadvantages for native children caused by lower parental education. However, with its pilot investments in targeting this subgroup of native children, it has pioneered a new approach in the Netherlands.

A quantile regression shows that the zero mean effect does not hide any impact in some point of the distribution. There are some indications, though, that child-centred ECE approaches produce positive effects for children from less educated parents, which are a target group of such early compensation investments. These results give a first indicative support to policymakers' plans to extend ECE programmes in South Limburg to children of less educated, native parents in more rural areas across the Netherlands. However, the results also raise questions about why no significant effects have been reached yet.

The lack of evident effects could be due to measurement errors, in particular in the child outcome and in the studied indicator for child-centred ECE experiences – normality might be absent in the dependent variable with a long bottom tail and truncation at the top. However, the studied child outcomes from Cito, an independent test provider, are less likely to contain measurement errors; they are applied regularly at most Dutch elementary schools and have been used in a large number of studies. Besides, the Dutch language test should be especially well suited to assessing the linguistic stimulation that child-centred ECE programmes provide. With respect to the studied ECE indicator, there is no obvious indication for measurement errors: I have cross-validated the information about ECE provision with an additional survey and secondary information, and excluded observations for which I cannot identify with full certainty whether or not the child attended a preschool that applied child-centred ECE.

There may be a number of other reasons why I cannot detect statistically significant effects on most of the estimated treatments. One potential reason is that the fraction of children attending child-centred ECE programmes is still very low, possibly below the minimum detectable impact. Correspondingly, the *Moelejaan* data I used confirms Leseman's (2002) argument that the intensity of early childhood care and education experiences tends to be insufficient to lead to sizable effects. And ECE programmes at preschool kindergartens and elementary schools in South Limburg were at an early stage of implementation at the time of data collection, which is in line with Driessen (2004); he suspects that too little professionalization of ECE provisions causes a lack of significant development gains.

Effects might be measurable in future, when more experience with child-centred ECE methods has been acquired by the preschool staff, ECE applications involve the parents more actively, and more children of the target group experience such

ECE. A large share of the children who could benefit from child-centred ECE preschool had not yet attended such preschools in the studied time period. Also, I assess test observations of children rather soon after their ECE experience at a preschool kindergarten. ECE programmes are usually continued in elementary schools, and thus some effects may show up only later on during the children's schooling.

Finally, Nap-Kolhoff et al. (2008) stress the lack of sufficient data about the quality of the child's early childhood care and education experience to provide informed evidence. The study at hand still has insufficient quality information, while effects estimates have to be averaged, e.g. across preschool groups and days, and child-centred ECE can only be compared to the joined alternative of teacher-based ECE and regular preschooling. Whereas this study uses new information about the type of ECE programme, it builds on the assumption that the application of an ECE programme reflects the full application of the quality conditions considered to be, in theory, a prerequisite for being effective. Even so, there is evidence that a majority of ECE providers do not yet fulfil all those conditions in practice.

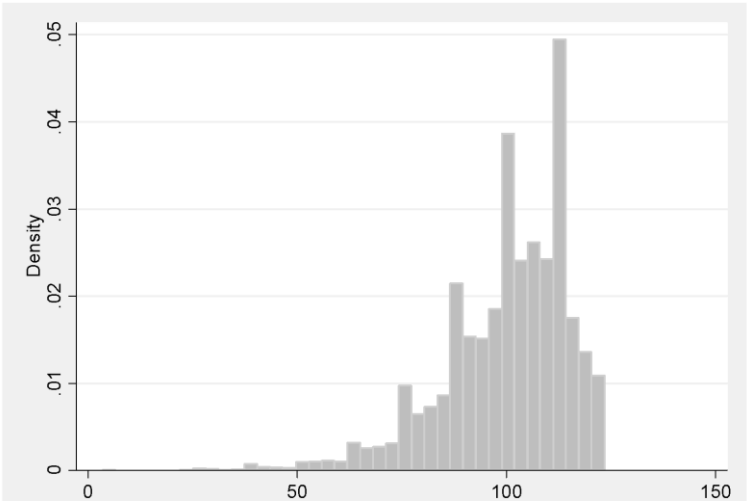
Further research on the quality of preschooling and its extension to bigger groups of disadvantaged groups is needed, which implies that municipalities introducing child-centred ECE programmes such as *Speelplezier* should consider devoting sufficient money to the collection of appropriate data to evaluate the effectiveness of the programme. The KAANS research centre at Maastricht University is working on such an extended data collection and research agenda for South Limburg. This research will also look more into the impact of child-centred ECE on children's socio-emotional and behavioural development. The strong interactive character of such child-centred ECE approaches suggests that positive effects may be found in those domains.

I conclude, in line with the literature, that various sample distortions impede finding significant and robust effects of investing additional funding into child-centred early childhood education programmes.

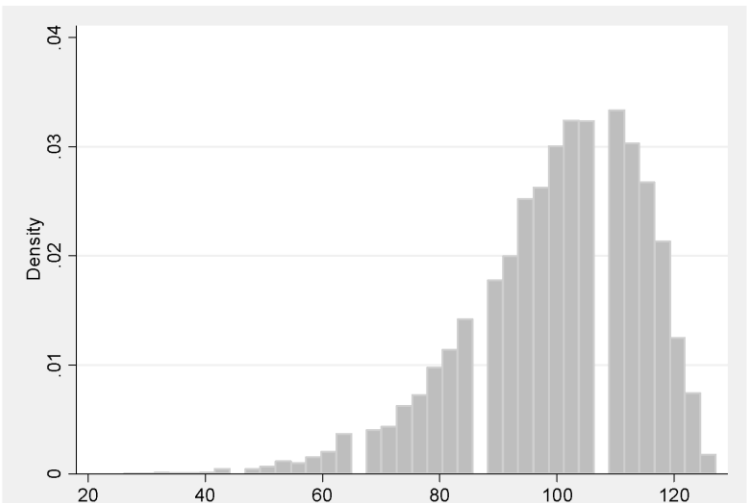
Appendix III

Figure III-1: Distribution of standardized child outcomes

a. Dutch language test, M2



b. Cognitive test, M2



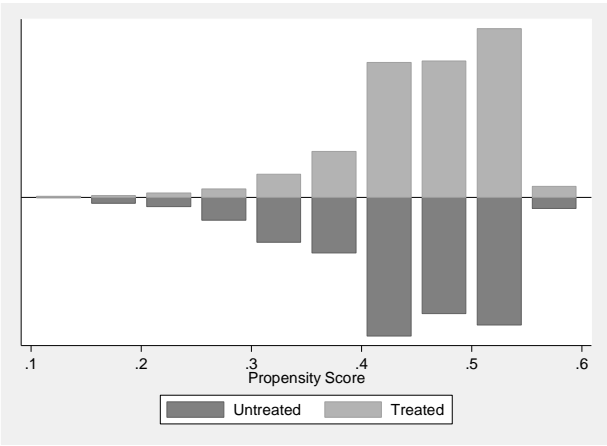
Note: *Cito* test scores have been standardized to a mean of 100 and standard deviation of 15.

Table III-5: Test scores in five categories and averages of standardized scores

Dutch language test, M2	Average standardized scores	Frequency	Cognitive test, M2	Average standardized scores	Frequency
22-41	77.90	172	15-28	76.97	161
42-45	94.43	229	29-31	91.92	214
46-48	101.78	245	32-34	100.63	318
49-50	107.56	233	35-36	107.83	224
51-55	114.63	429	37-41	116.31	392

Figure III-2: Calculated weights

- a. Overlap between population (*untreated*) and matched sub-sample that is studied (*treated*)



- b. Distribution of calculated non-response weight

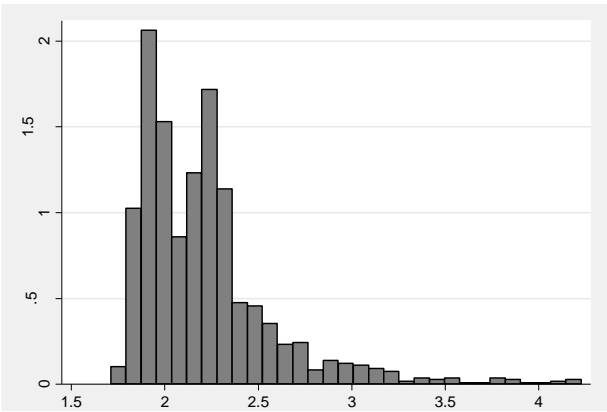


Table III-6: Descriptive statistics

Control variable	N	Mean	Std. Dev.	Min.	Max.
Child-centred ECE (attendance dummy)	1321	0.09	0.29	0.0	1.0
Total length of preschool attendance (in units of 10 half-day sessions)	1321	14.54	8.82	0.0	72.0
Total length of daycare attendance (in units of 10 half-day sessions)	1321	4.52	11.64	0.0	90.0
Total length of childcare at home (in units of 10 half-day sessions)	1321	73.36	33.00	0.0	180.0
Dutch language test (taal voor kleuters, M2-testing)	1321	102.37	13.14	37.8	123.5
Cognitive test (ordenen, M2-testing)	1321	102.07	13.74	38.0	127.1
Gender (boy = 1)	1321	0.51	0.50	0.0	1.0
Age at time of testing (in years, middle of 2 nd grade)	1321	5.77	0.38	3.7	7.3
Ethnic minority (min. one parent born abroad = 1)	1321	0.13	0.34	0.0	1.0
Parenthood (single parents = 1)	1321	0.05	0.22	0.0	1.0
Number of children in household	1321	2.12	0.71	1.0	7.0
Mother's unemployment	1321	0.02	0.15	0.0	1.0
Highest parental education: level 1 (no / elem. / lower educ.)	1321	0.13	0.34	0.0	1.0
Highest parental education: level 2 (middle / professional educ.)	1321	0.52	0.50	0.0	1.0
Highest parental education: level 3 (higher professional / acad. educ.)	1321	0.35	0.48	0.0	1.0
Number of children's books at home (1 unit = 5 books)	1321	4.46	0.95	1.0	5.0
Father's Dutch language skills (5 point Likert scale)	1321	4.50	0.69	2.0	5.0
Net monthly hh-income (logarithmic scale)	1321	1.66	0.43	0.0	2.4

Note: These figures are based on the studied sub-sample of the *Moelejaan* data for South Limburg; non-response weights are applied (see section 5.1).

Table III-7: Length of different childcare experiences

	Obs.	Mean	Std. Dev.	Parental care	Family care	Nanny care	Home care	Daycare	Preschool	Total reported childcare
Parental care (home-based)	1321	52.10	32.12	1.00						
Family care (home-based)	1321	18.04	19.85	-0.31	1.00					
Nanny care (home-based)	1321	3.72	12.06	-0.13	-0.04	1.00				
Overall home-based care	1321	73.87	32.56	0.75	0.29	0.22	1.00			
Daycare (centre-based)	1321	4.49	11.69	-0.15	-0.13	0.01	-0.23	1.00		
Preschool kindergarten (centre-based)	1321	14.46	8.73	-0.02	0.08	-0.01	0.02	-0.15	1.00	
Total childcare	1321	92.82	32.91	0.68	0.27	0.21	0.92	0.09	0.24	1.00

Note: These unweighted figures (in units of 10 half-day sessions) are based on the sub-sample covered by the *Moelejaan* survey for South Limburg for which sufficient information has been available. In this survey, we have asked parents about the types of childcare used before enrolling their child at elementary school, including questions about the number of half-day sessions per week and years of total childcare attendance. Most children enrol at school when they are about four years old. Children can potentially experience a maximum of about 1400 half-days of childcare (= 3.5 years before enrolment x 2 half-day sessions x 5 working days per week x 40 weeks per year excluding holidays). Descriptive statistics of the collected data are approximately within that range. Still, the average total number of reported childcare half-days (=927) is smaller than this total possible number of half-days; some parents do not provide complete information on all possible answer options, and other possible childcare alternatives have not been asked.

Table III-8: Appropriate parental quoting of *Speelplezier* ECE method

The ECE method (*Speelplezier*) quoted by parents in the survey resembles the truly applied ECE method that has been linked via the preschool's name

	Not appropriate	Appropriate
Highest parental education: level 1 (<i>no / basic / lower educ.</i>)	65.1%	34.9%
Highest parental education: level 2 (<i>middle / professional educ.</i>)	67.8%	32.2%
Highest parental education: level 3 (<i>higher professional / acad. educ.</i>)	65.0%	35.0%
Total	67.0%	33.0%
Observations (<i>within full sample</i>)	1173	578

Table III-9: Univariate sample comparison, with and without ECE information

Variable	N ₁ (<i>studied sub-sample</i>)	N ₂ (<i>non-studied sample</i>)	mean comparison	t-test at 5% significance
Gender (<i>boy</i> = 1)	5693	10645	mean ₁ < mean ₂	(marginally) significant difference
Age at time of testing (<i>in years, middle of 2nd grade</i>)	2634	4820	mean ₁ < mean ₂	significant difference
Ethnic minority (<i>min. one parent born abroad</i> = 1)	5692	5283	mean ₁ < mean ₂	significant difference
Parenthood (<i>single parents</i> = 1)	5509	4897	mean ₁ < mean ₂	significant difference
Number of children in household	5684	5171	mean ₁ > mean ₂	no significant difference
Mother's unemployment	5392	5051	mean ₁ < mean ₂	significant difference
Highest parental education (<i>3 attainment levels</i>)	4685	8806	mean ₁ > mean ₂	significant difference
Number of children's books at home (<i>1 unit</i> = 5 books)	5502	5003	mean ₁ > mean ₂	significant difference
Father's Dutch language skills (<i>5 point Likert scale</i>)	4850	5490	mean ₁ > mean ₂	significant difference
Net monthly hh-income (<i>logarithmic scale</i>)	4527	4232	mean ₁ > mean ₂	significant difference

Table III-10: Multivariate probit regression for different samples

Variable	Coefficient	Standard Error	Z	P
Gender (<i>boy = 1</i>)	-0.04	0.05	-0.93	0.35
Age at time of testing (<i>in years, middle of 2nd grade</i>)	0.06	0.06	1.05	0.29
Ethnic minority (<i>min. one parent born abroad = 1</i>)	-0.23	0.07	-3.32	0.00
Parenthood (<i>single parents = 1</i>)	-0.16	0.11	-1.50	0.13
Number of children in household	0.02	0.03	0.77	0.44
Mother's unemployment	-0.31	0.13	-2.31	0.02
Highest parental education: level 2 (<i>middle / professional educ.</i>)	0.19	0.07	2.56	0.01
Highest parental education: level 3 (<i>higher professional / acad. educ.</i>)	-0.01	0.08	-0.14	0.89
Number of children's books at home (<i>1 unit = 5 books</i>)	0.05	0.02	2.05	0.04
Father's Dutch language skills (<i>5 point Likert scale</i>)	-0.04	0.03	-1.21	0.23
Net monthly hh-income (<i>logarithmic scale</i>)	0.05	0.06	0.73	0.47
Constant	-0.49	0.40	-1.22	0.22

Note: Multivariate probit regression for different sub-samples (with and without ECE information).

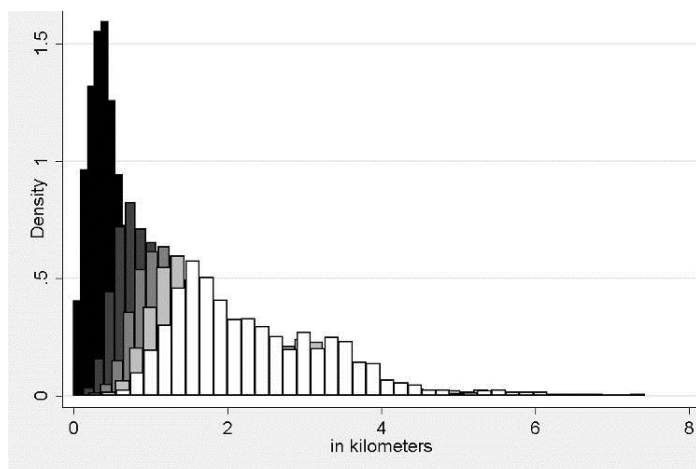
Table III-11: Preschool comparison, with and without child-centred ECE

Variable	N ₁ (child-centred ECE)	N ₂ (other ECE or regular)	mean comparison	t-test at 5% significance
Ethnic minority (min. one parent born abroad = 1)	634	5058	mean ₁ < mean ₂	no significant difference
Parenthood (single parents = 1)	609	4900	mean ₁ < mean ₂	no significant difference
Number of children in household	643	5041	mean ₁ > mean ₂	no significant difference
Mother's unemployment	512	4880	mean ₁ > mean ₂	no significant difference
Highest parental education (3 attainment levels)	548	4137	mean ₁ > mean ₂	(Marginally) significant difference
Number of children's books at home (1 unit = 5 books)	571	4931	mean ₁ < mean ₂	significant difference
Father's Dutch language skills (5 point Likert scale)	620	4870	mean ₁ < mean ₂	significant difference
Net monthly hh-income (logarithmic scale)	494	4033	mean ₁ < mean ₂	significant difference

Table III-12: Distance information between homes and preschools

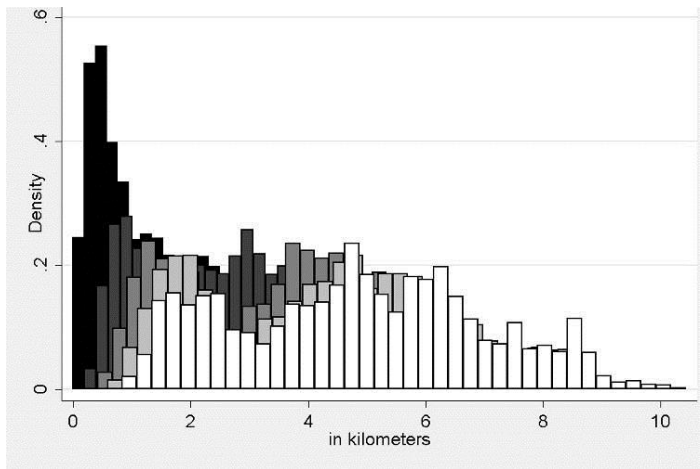
	Obs.	Mean	Std. Dev.
Distance to chosen preschool	1309	1.52	2.99
Distance to closest preschool	1309	0.59	0.42
Distance to closest child-centred ECE preschool	1309	1.90	1.58
Distance of chosen preschool compared to average 5 closest alternatives	1309	0.99	2.18
Distance of chosen preschool compared to average 5 closest alternative options of child-centred ECE preschools	1309	0.50	1.17
Distance of closest child-centred ECE preschool over closest alternative option	1309	218	3472

Figure III-3: Distances to closest preschools without child-centred ECE



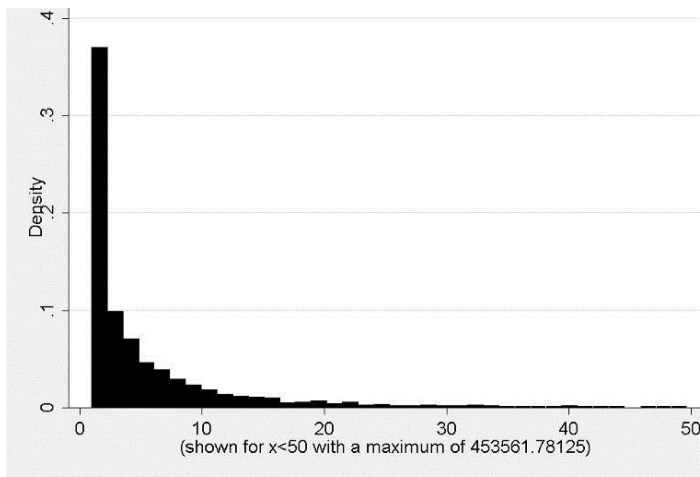
Note: Shading goes from the closest available preschool in black bars to the 5th closest in white bars.

Figure III-4: Distances to closest preschools with child-centred ECE



Note: Shading goes from the closest available preschool in black bars to the 5th closest in white bars.

Figure III-5: Closest preschool with child-centred ECE over alternatives



Note: Relative distance of closest preschool with child-centred ECE over any alternative preschool.

Data Annotation

The dataset used for the third study was collected to evaluate the *Moelejaan* policy pilot that aimed to extend early childhood education applications to all preschool kindergartens in smaller municipalities of South Limburg in 2009-2010 (Cox, Jungbluth, & Rodigas, 2008). A large-scale data collection was carried out in 2008/09 to produce a baseline image of the whole region, including the bigger municipalities that had started earlier in making early childhood education investment. The data collection covered more than 200 elementary schools in the region and thus about the whole child population of the region (Jungbluth, 2009).

The KAANS Research Centre of Maastricht University (formerly called SJOANS) took the lead in this data collection (see also <http://www.kaans.nl>). Being interested in using the baseline dataset of the *Moelejaan* pilot for my dissertation, I got the opportunity to be involved in the data collection from the very beginning. An initial study of 16 schools in the municipality of Maastricht that I carried out in late 2008 allowed me to draft and pre-test the main data collection instruments, i.e. procedures to tab basic registry information and test scores from school administrations and survey questionnaires for parents and headmasters of the schools as well as cooperating daycare centres and preschool kindergartens (Bauchmüller & Jungbluth, 2008).

The development of those instruments was inspired by the national PRIMA cohort study, which I used for the first study (chapter II). Together with colleagues of the KAANS Research Centre, I developed the initial instruments further, so that they could be used in 2009 for the full-scale *Moelejaan* data collection in the whole region, and provided further guidance on the fieldwork that was taken over by the regional education service provider Consent, which further automatised many of the procedures (Rodigas, 2009). In the meantime, I collected additional secondary information to complement the *Moelejaan* data, e.g. on educational policies of municipalities in the region, individual school statistics from national public records, and socio-economic figures on municipalities from national statistics.

On behalf of KAANS, I checked incoming data across the different sources and oversaw that they can be successfully integrated into a larger comprehensive dataset via anonymized identifiers. When the fieldwork of the baseline data collection was concluded in late 2009, I handed over the administered and tested datasets to the KAANS Research Centre, which has added additional records in the meantime; it strives to allow a comprehensive analysis of children in Limburg and their education and performance at various ages.

The KAANS Research Centre applies the highest professional standards when dealing with sensitive personal and confidential data, both with respect to anonymisation and secure handling. KAANS deletes data of individual children

from the datasets if their parents request it - they had the opportunity to do so, for instance, in the survey questionnaire. Organizations who are involved in the data collection and researchers who make use of the data need to sign a declaration of consent with the centre's use of professional standards in dealing with sensitive information. The Dutch Data Protection Authority has been informed about the data collection. Observations of children, households and preschool kindergartens have been made fully anonymous before any analysis took place.

Chapter IV - Long-run benefits from universal high-quality preschooling (*study III/III*)

1. Introduction³⁹

Social science research increasingly acknowledges the importance of early childhood development. Early investments lead to high private and social returns (Cunha et al., 2006). Esping-Andersen (2006, 2008) argues that the assurance of high-quality daycare could be the single most effective policy for reducing inequalities in education and income.

While a number of evaluation studies document short-term effects of early childhood care (Datta Gupta & Simonsen, 2010; Jensen, 2009), only a few studies of long-term effects of childcare exist. Moreover, most previous studies either analyse the effects of daycare *per se* or compare different *types* of daycare, e.g. locally run preschool versus private childcare. More knowledge of how the *quality* of daycare affects child outcomes, especially for disadvantaged children, is needed. Little is yet known about which preschool quality characteristics are most beneficial for children or about whether disadvantaged children particularly benefit from high-quality preschools.

This chapter analyses whether variation in the quality of preschools is related to long-term child outcomes. Our approach is inspired by recent advances in the literature on the economics of education, which focuses on how individual schools and teachers contribute to cognitive and intellectual development (cf. Dronkers, 2010; Dustmann, Puhani, & Schönberg, 2012; Heckman, 2008; Kramarz, Machin, & Ouazad, 2009; Rivkin, Hanushek, & Kain, 2005). Like Havnes & Mogstad (2010) and Datta Gupta & Simonsen (2010), we examine a universal daycare system in the Scandinavian tradition. To investigate the equalization potential of preschool attendance, we study the heterogeneity of quality effects across several subgroups of vulnerable children.

Using an extensive register data set with information on child outcomes, children's preschool affiliation, and the staff in each preschool institution, we generate five different indicators of quality: (1) the staff-to-child ratio, (2) the share of male staff in the preschool, (3) the share of trained teacher staff in the preschool, (4) the share of ethnic minority staff, and (5) the stability of the staff (inverse staff turnover). As previous research has shown that test scores are strong predictors for future

³⁹ Earlier versions of this chapter have been released as AKF Working paper, 2011(2), and as CSER Working Paper, 2011(8).

educational and labour market outcomes (Connolly, Micklewright, & Nickell, 1992; Currie & Thomas, 2001), we measure child outcomes by grades in written Danish obtained from the 9th grade final exam at the end of primary school.⁴⁰ Using Ordinary Least Squares (OLS), we ask two policy-relevant questions: *Are individual 9th grade test results correlated with the quality of children's preschools? Can we establish a causal relationship between the quality of preschool and child outcomes?* To ascertain the causality of effects of preschool quality on child outcomes, we do a sensitivity analysis using instrumental variable (IV) analysis. As suitable instruments that capture parents' potential selection of preschools are difficult to find, we test eight different instruments in our specifications.

Our main contribution lies in establishing modest but long-term effects of preschool quality. Thus we extend the literature on preschool arrangements by examining whether variation in quality aspects can explain some of the variation in later child outcomes. As the Scandinavian countries have been the forerunners in universal early childcare, results for Denmark are likely to have wider applicability throughout the advanced economies.

The chapter is organized as follows: Section 2 reviews the literature on early childhood interventions. Section 3 describes the data and institutional framework of preschooling in Denmark. Section 4 elaborates on our quality indicators. Section 5 describes the evaluation methodology. Section 6 presents the results of our empirical analysis, and Section 7 presents conclusions.

2. Background and related literature

Recent research emphasizes the importance of early childhood development for long-term outcomes across a wide range of development domains (Heckman, Stixrud, & Urzua, 2006). Measures of long-term outcomes include high-school enrolment, crime rates, health, early pregnancies, job aspirations, labour-market success, career outcomes, and life expectancy. Early childhood is considered a sensitive period for brain development and language acquisition, and early learning is crucial for later learning (Heckman et al., 2002; Knudsen et al., 2006). Considerable heterogeneity in child outcomes exists in early childhood, much of it attributable to environmental factors such as family background and parental taste (Heckman, 2008; Mayer, 1997). The participation of disadvantaged children in targeted high-quality childcare programmes can be more effective than giving their parents unrestricted cash transfers (Currie, 1998).

⁴⁰ Primary school in Denmark is K-9. However, because kindergarten starts at age 6, 9th graders are aged 15-16 when they finish primary school.

Previous evaluations of long-term effects of daycare arrangements have focused on small-scale, randomized controlled trial programmes targeted at disadvantaged children, e.g. the Perry Preschool programme (Schweinhart & Weikart, 1981) and the Abecedarian programme (Campbell et al., 2008). While some of these studies suggest that gains from preschool eventually fade, especially if continual and sufficient support for children at risk is not assured (Esping-Andersen et al., 2011; Reynolds, 1993, 2000), most studies suggest lasting effects of daycare (see, for example, Barnett, 1992, 1995, 2011; Melhuish, 2011; Nores & Barnett, 2009; Vandell et al., 2010).

While the literature on the introduction or expansion of universal daycare finds some mixed results, most studies find positive results. A number of Head Start evaluations find positive effects, especially for disadvantaged children (Currie, 2000; Deming, 2009; Ludwig & Miller, 2007). Similar results appear in a number of European evaluations, e.g. Felfe & Lalive (2010), Dumas & Lefranc (2010), and Havnes & Mogstad (2010, 2011). However, a study by Baker, Gruber & Milligan (2008) finds significant negative effects on behavioural and motor-social skills from the transition to a regime with large-scale, highly subsidized childcare.

A few studies compare differences between *types* of care, e.g. formal centre-based care vs. informal private care and care at home (Bernal & Keane, 2006; Datta Gupta & Simonsen, 2010; Esping-Andersen et al., 2011; Gregg, Washbrook, & ALSPAC Study team, 2003; Waldfogel et al., 2002). For example, using data from the U.S. National Longitudinal Survey of Youth, Bernal & Keane (2006) find that children benefit from being enrolled in formal care (centre-based care and preschool), but experience adverse outcomes when participating in less expensive informal care. Datta Gupta & Simonsen (2010) find no significant differences in non-cognitive skills for 7-year-olds when comparing preschool to home care. However, for boys of less educated mothers, private daycare appears negatively related to non-cognitive abilities. Esping-Andersen et al. (2011) compare three types of care across Denmark and the United States: high-quality centre-based care, lower-quality child-minding, and family care. For Denmark, they find a significant and positive effect of high-quality care on reading skills that lasts until age 11, especially for disadvantaged children. For the U.S., while the effects are not lasting, formal school- or centre-based care has significant positive effects at school entry. They argue that selection into unequal quality of schooling in the U.S. may erode early development gains, whereas the more equal Danish system may manage better to preserve early gains.

A number of studies quantify the importance of the quality of universal daycare offered. Lasser & Fite (2011) argue in favour of universal high-quality preschool provisions, stressing the importance of teacher preparation. Moreover, Gregg et al. (2005) find no adverse effects on child development from daycare, suggesting that

high-quality, centre-based care neutralizes potentially negative effects of maternal employment. In a recent study of Project Star, Chetty et al. (2011) find that smaller preschool class sizes, a longer teacher experience in kindergarten, and a stronger peer group are positively associated with, for example, better school attainment and higher earnings later in life. In contrast, Blau (1999) suggests that daycare characteristics have little association with child development.

3. Data and institutional framework

3.1. Data

We use a unique longitudinal dataset based on administrative register data created by Statistics Denmark. This dataset includes data from the daycare register, which covers about 95% of the municipalities. Statistics Denmark has successfully matched more than 2000 preschools to workplaces from the Integrated Database of the Labour Market (IDA), thus linking children and staff in preschools (cf. Gørtz & Andersson, 2010). Our sample consists of an almost complete cohort of children who finished primary school's 9th grade in 2008; these children were born around 1992 and were typically 15-16 years old at the end of 9th grade. Our outcome variable is children's grades from final exams in written Danish. These tests are nationally comparable and externally evaluated.⁴¹

As the registers provide comprehensive longitudinal information on each citizen's origin and socioeconomic background, we generate variables for the children's family characteristics, e.g. parental education, household income and number of siblings in 1998 (the last year before this cohort entered primary school). Measures of preschool quality (see section 4) are constructed from register information on preschool teachers. Our study focuses on the effect of a variation in preschool quality for the sample of children who attended preschool.

3.2. Organization of daycare and preschool in Denmark

The majority of Danish daycare facilities are organized and operated by the municipalities. Municipalities provide centre-based early daycare at nursery centres for the 0-3-year-olds, preschools for children aged 3-6, age-integrated institutions (for the 0-6-year-olds) with a separate preschool compartment, and after-school programmes for school children. In addition, municipalities can

⁴¹ We also estimate effects on a number of other child outcome variables, i.e. tests of written English skills, mathematic problem-solving skills and Danish verbal skills. However, given that the preschool period probably is most sensitive for language development, we focus our discussion on written Danish tests.

organize family daycare where young children (usually 0-3-year-olds) are cared for by a childcare provider, usually in her or his private home. This chapter focuses on centre-based preschools and the preschool sections of age-integrated institutions.

Almost half of all institutions are independent, i.e. semi-private non-profit organizations that basically function as municipal institutions but with more autonomy from the municipality and more parental influence.⁴² While parental payment is the same for institutions within a municipality, fees vary across municipalities. Parents usually pay 20-30% of the actual costs of a preschool slot. The municipality can decide how much of the cost parents should bear, but national law sets an upper cap.

While parents cannot freely choose among institutions, they can submit an application to the municipal administration with a prioritized list of preferred institutions. The municipality then adds the children to a waiting list according to their date of birth. While municipalities allocate children to daycare institutions, the centres are not directly involved in the allocation process. As excess supply is costly, municipalities try to distribute children so as to fill all daycare institutions. Waiting list position determines probability of access. If parents limit their choice to a specific institution and thus a specific preschool quality, they usually have to wait longer for an opening. Children with specific needs or children at risk may jump the waiting list. Moreover, enrolling younger siblings in the same institution as their older sibling is a priority. Nevertheless, the municipal authority makes the final decision, with parental preferences being merely a guide in the process.

Guaranteed Access to Pre-School from the age of 6 months (GAPS) has been part of government policy since the mid-1990s. This policy has been transmitted to municipalities at a varying pace; about 72% of the children in our sample attended preschools in municipalities with GAPS. The implementation of GAPS coincided with a substantial expansion in daycare, particularly for children under the age of 3.

3.3. Attendance rates and parental selection

Danish parents use daycare on a large scale. In 1998, almost 70% of 1-year-old children attended some kind of formal daycare arrangement, and more than 90% of 3-6-year-olds attended daycare (see Table IV-1). Attendance rates have continued to rise over the last decade.

⁴² Legislation on educational quality and provisional conditions for hygiene and safety are the same for both types, differences in observed quality are numerically small but significantly different from 0. Nonetheless, differences in outcomes across types are insignificant.

Table IV-1: Child and preschool population in 1998

	Preschool kindergarten		Age-integrated institution		Total child population	Total attendance rate ^a
3 years	30,019	42.3%	20,071	28.3%	71,000	86.9%
4 years	39,146	57.0%	22,025	32.1%	68,686	90.3%
5 years	40,720	58.8%	21,334	30.8%	69,281	93.1%
6 years ^b	7,873	11.9%	9,679	14.6%	66,139	84.8%

Source: (Statistics Denmark, 1998; figures per 01.01.1998).

^a Attendance figures for all types of daycare, including preschool and age-integrated institution. Calculations based on register data.

^b Children usually enrol in primary school the year they turn 6; preschool statistics are collected later in the calendar year.

Children's sorting into preschools is a consequence of a number of prior parental choices. First, parents choose a municipality of residence according to various characteristics such as tax rates, infrastructure, and local labour markets. Second, within municipalities, parents choose to settle in a particular neighbourhood. This choice is more likely to be governed by information on local public schools than that on local preschools. Third, given the neighbourhood, parents prioritize among daycare institutions in the neighbourhood, according to perceived quality, feedback from neighbours and friends, experiences of siblings, or the parents' educational values and priorities. However, given long waiting lists, parents have no guarantee of obtaining their preference, and the final allocation of children into institutions is done by the municipality. Hence, as children's sorting into preschools seems to be exogenous, we can interpret a possible correlation between preschool characteristics and child outcomes as causal.

Consequently, combining these characteristics of municipal allocation procedures with our assumptions about parental choices and their implications for children's sorting into childcare institutions, we argue that allocation of children into preschools is largely exogenous. Nevertheless, to deal with possible endogeneity, we test a number of instrumental variables.

4. Preschool quality

Behrman & Birdsall (1983) stress that looking solely at the quantity of schooling is misleading because the variation in quality is substantial. The same risk applies to preschooling. Classic quantity measures are the duration in daycare as measured by the number of years, and weekly hours of attendance as provided in surveys or administrative records (Datta Gupta & Simonsen, 2010).

Despite the high average quality of preschools, considerable variation exists in the resources spent on preschools across, as well as within, municipalities. A number of structural aspects are important for assessing the quality of a preschool; these include type of care, staff qualifications, group size, physical facilities, quality of management, easy accessibility, peer group composition, the educational approach, the share of male teachers, teacher turnover, and parental involvement (Gørtz & Andersson, 2010; Marshall, 2004).

From the register data, we therefore generate a set of quality characteristics for preschools in 1998.⁴³ For each preschool, we calculate five indicators of quality that we expect to have positive impacts on child outcomes: (1) the staff-per-child ratio, (2) the share of male staff, (3) the share of trained preschool teachers among the staff, (4) the share of ethnic minority staff, and (5) the stability of staff per institution compared to the previous year.

Underlying our choice of the first indicator, the staff-per-child ratio, is the hypothesis that increasing the number of staff members per child results in more staff time being spent with each child, improving the child-teacher interaction (see, e.g. Frede, 1998). More time implies more attention to the child's individual needs and development, following arguments in the literature (cf. Card & Krueger, 1996) on class size.⁴⁴ Evidence for positive effects of smaller class size is found by Fredriksson & Öckert (2008) who study a drastic change in the availability of school resources caused by a policy reform in Sweden in the early 1990s. In their meta-analysis of 123 comparative studies of early childhood interventions, Camilli et al. (2010) find support for positive effects of small-group, teacher-directed instructions (see also Copple & Bredekamp, 2009).

Often a mix of trained and untrained staff is directly involved in childcare. Trained preschool teachers are those who have completed formal training in preschool education. Untrained staff is not required to have any such education. Our first indicator includes both trained and untrained staff.

⁴³ About 1.6% of the preschools and 7.5% of the age-integrated institutions were reported at the municipal instead of institutional level. For those municipalities, preschool quality characteristics were generated at the municipal level, and excluding those observations does not change estimates noticeably. For 9.3% of municipalities, daycare registers are not available for the studied period.

⁴⁴ If children with special needs (unobserved in the data) are placed in institutions with more staff per child, there may be an inference problem, as our estimates of the impact of the staff-to-child ratio may then be downward biased. We take this problem into account in the IV estimations using county averages as an instrument.

Our second indicator, the share of male staff, derives from a recent discussion linking mounting school problems of boys to the lack of male role models in early childhood (see, for example, Dee, 2005a, 2005b, 2006). This literature argues that boys need a specific type of care that presumably only male staff can provide. However, most preschool teachers are women.⁴⁵

Our third quality indicator, the share of trained staff, derives from the assumption that educated staff is better equipped both to understand the individual needs of a child and to emphasize learning and school preparation. Municipalities have targets for these percentages. Braga, Checchi & Meschi (2011) cite Hanushek (2010) when stressing the importance of teacher qualifications for child development. However, they find evidence that such qualifications are likely to be more and more important as children proceed up the educational ladder.

Our fourth quality indicator reflects the share of ethnic minority staff. Dee (2004, 2005a) suggests two theoretical arguments for potentially positive effects of ethnic minority staff: active teacher biases and passive student responses. Active teacher biases means that a teacher's behaviour towards a child will likely be influenced by the teacher's perception of that child. The behaviour of a Danish-"majority" teacher towards a minority child could be biased by stereotypes about that minority group, while minority teachers – with a better understanding of minority cultural experiences – might be better for taking care of these children. The theory of passive student responses means that a minority child sees a minority teacher either as a role model or a good example of integration with the majority group. This theory argues that a more culturally and ethnically diverse staff is likely to better prepare children for a more diverse society and potentially help minority children. Other arguments stress the importance of having caregivers who speak the child's native language and of recruiting ethnic minority teachers for relieving "difficult-to-staff" or "high-attrition" schools (Clewett & Villegas, 1998). Despite some empirical evidence of the expected positive effects of a culturally more diverse teaching staff at schools (Villegas & Jordan Irvine, 2010), little is yet known about whether these arguments also hold for preschool staff.

Our fifth indicator, staff stability (inversely proportional to staff turnover), reflects the share of staff that stay employed in the same institution from one year to the next. If a large share of the staff changes at a centre from year to year, there is less retention of centre- and child-specific knowledge, possibly reducing attention to the needs of individual children. Furthermore, a higher turnover rate demands the

⁴⁵ In 1997, a Danish Governmental Child Committee proposed to further the recruitment of men in child and youth educator study programs, teacher study programs, and daycare facilities (cited in: OECD, 2000).

investment of more resources in the hiring process. Conversely, stability of human capital accumulated in a preschool may lead to better child outcomes.⁴⁶ Ronfeld et al. (2011), for example, find that a higher teacher turnover for 4th and 5th graders has a detrimental influence on student achievement.

Table IV-2: Quality characteristics

Preschool kindergarten	Obs.	Mean	Std. dev.	Min.	Max.
Number of staff members per child ^a	20,345	0.228	0.065	0.071	0.727
Share of male staff members	20,345	0.114	0.103	0.000	0.667
Share of trained preschool teachers	20,345	0.462	0.125	0.000	1.000
Share of ethnic minority staff members	20,345	0.038	0.070	0.000	1.000
Staff stability ^b	20,345	0.677	0.140	0.150	1.000
Age-integrated institution					
Number of staff members per child ^c	10,099	0.228	0.065	0.082	0.605
Share of male staff members	10,099	0.134	0.106	0.000	0.571
Share of trained staff	10,099	0.436	0.110	0.000	1.000
Share of ethnic minority staff members	10,099	0.052	0.073	0.000	0.714
Staff stability ^b	10,099	0.654	0.129	0.182	1.000

a A small number of outliers in staff-per-child ratios, values below the 0.5% and above the 99.5% percentile, are recoded as missing.

b Staff stability is defined as the share of staff staying employed in the preschool from one year to the next. Calculation is based on institutions that were in the dataset in both 1997 and 1998. Missing values have been replaced by averages.

c These values are reweighted towards the same averages as those for preschools. The original mean was 0.265, and the original standard deviation was 0.082.

Table IV-2 gives summary statistics of our five indicators of quality. On average, one staff member takes care of about five children. The staff indicator includes not only the trained preschool teachers but also the assistant teachers and janitorial staff. Overall, although one in ten staff members is male, and about one in two staff members is a trained preschool teacher, the variation is substantial. For both preschool and age-integrated institutions, about one in three employees changes

⁴⁶ High turnover rates do not necessarily imply that most of the staff has changed but rather that some staff positions are subject to relatively frequent changes.

schools every year. However, the average quality did not significantly change from 1996 to 1998, and quality factors are only very weakly correlated.⁴⁷

The indicators are normally distributed, except for the shares of male and ethnic minority staff, whose shares are skewed to the left on a 0-to-1 scale. For age-integrated institutions, we need to impute indicators for the part of the institution that is directed towards 3-6-year-olds. Age-integrated institutions usually separate children by age groups and have, for the 3-6-year-olds, staffing levels similar to those of preschools. We therefore reweigh staff-per-child ratios to the same overall average and variance for staff-per-child ratios as those for preschools. We hypothesize that higher values for the indicators lead to positive development gains for children.

5. Methodology

Our empirical model rests on a production function for child development. This approach finds its inspiration in work by Heckman (2008), Ruhm (2005) and Todd & Wolpin (2003). We model the production of child i 's outcome (CO_i) as a linear function of individual characteristics of the child, including the length of the child's preschool history (X_i), family characteristics assumed to account for part of the variation in children's initial development (F_i)⁴⁸, quality characteristics of the preschool institution of child i (I_i), and contextual measures of the preschool (P_i). Moreover, the child outcome is determined by an average baseline score β_0 and a general idiosyncratic error term as shown in equation (IV.1):

$$CO_i = \beta_0 + \beta_1 X_i + \beta_2 F_i + \beta_3 I_i + \beta_4 P_i + \varepsilon_i \quad (IV.1)$$

Child characteristics, X , include gender (a dummy for being male), the child's age (in the preschool period 1998), and a dummy for being an ethnic Dane. Family characteristics, F , include a dummy for living in a single-parent household, the number of children in the household, the mother's age and level of unemployment, the highest educational level of the parents, and a logarithmic term of gross income per household, all measured during the child's preschool period in 1998.⁴⁹ The set

⁴⁷ The five indicators are weakly correlated with age and work experience. Centres with more ethnic minority or male staff tend to have somewhat younger staff.

⁴⁸ No initial child development outcome is available in the data.

⁴⁹ As father's unemployment has a large number of outliers, we do not use this variable. / As parental income may potentially be endogenously related to preschool attendance or the quality of preschool attended, many studies do not use income measured in the preschool period. To test for such endogeneity, we have done a sensitivity study with the Danish study showing that there are no significant changes in income across early and later childhood.

of peer indicators, P , reflects each preschool's share of boys, ethnic minority children, children with single parents, children with less educated parents (not more than primary schooling), and children from poor households. An overview of descriptive statistics on the control variables appears in Table IV-6.

This conditioning set covers the main factors identified in the literature as relevant determinants of child outcomes, through the direct channel of covering aspects of inherited abilities and the indirect channel of accounting for parental preschool choices. Björklund & Salvanes (2011), for example, confirm the importance of various family background factors for child development, stressing in particular the importance of parental education levels as well as the socioeconomic resources in the family. Early childhood is recognized as a highly sensitive period with respect to a child's resource endowments (Haveman & Wolfe, 1995). The literature attributes later inequalities in child outcomes to unequal access to such resources and proposes compensatory public investments, for example, by providing access to high-quality preschooling (Bowles et al., 2005; Esping-Andersen, 2006).

Family background factors also have an indirect effect on child outcomes through their impact on parental demand for high-quality preschool investments. Parents who are better endowed in terms of socioeconomic resources and cultural capital are more likely to demand a specific preschool arrangement. Parents base their choice with particular attention to the peer group at a specific preschool, expecting positive spill-over effects from a better group composition. However, the empirical evidence on such direct impact of a peer group on child development is rather weak (Gibbons, Machin, & Silva, 2006).

The analysis is based on a cohort graduating from Danish elementary schools in 2008. Due to left censoring in the daycare registers, we cannot observe whether children attended any other daycare institutions before 1995, i.e. before the child was around 3 years old. However, we account for the length of the preschool attendance as of 1995, meaning that we have uncensored, full preschool spells for the ages 3-6 years.⁵⁰ To allow for possible non-linearity, e.g. due to decreasing returns to preschooling, we also add a squared term of the length of the last preschool spell.

The empirical analysis focuses on grades from a compulsory final examination in written Danish by the end of 9th grade (equivalent to finishing primary school). Children typically take the exam when they are aged 15-16. Hence, we study effects of preschool about ten years after attendance. We consider the school trajectory to partly constitute a continuation of the initial preschool experience.

⁵⁰ On average, children attend preschool for 2.75 years and age-integrated institutions for 3.06 years.

Thus, the test outcomes after primary schooling show the full effect of preschooling. Since the 2006/2007 school year, the Danish grading system has used a scale that ranges from -3 (unacceptable) to +12 (excellent), with a score higher than 0 being a pass. For easier comparison with the international literature, we have standardized grades to a mean of 100 and a standardized deviation of 15. Table IV-7 compares the original and standardized scores.

To investigate whether high preschool quality is particularly beneficial for children at risk, we include a number of interaction terms between the quality measures and subgroup indicators. We study the following subgroups: (1) ethnic minority children, (2) boys, (3) children of less-educated parents (a maximum of primary education), (4) children of single parents, and (5) children living in households in the bottom quintile of gross incomes (henceforth, “poor” households). For simplicity we include only significant interaction coefficients in the results tables.

For the choice of the appropriate estimation technique, the crucial question is whether the quality of the preschool institution (I_i) is exogenous to the child outcome in equation (IV.1). In section III.C, we described the allocation of daycare slots administered by the municipalities. Parents have only limited influence on the final allocation of their child. Hence, given that I_i is to a large extent exogenous, OLS is consistent and efficient. However, to the extent that I_i is endogenous, OLS is inconsistent.

Endogeneity may occur if the preschool institution is not randomly assigned, e.g. if parents can influence the choice of preschool institution (see also Duncan et al., 2004). We assume that parents prefer high-quality preschooling, and if high-ability children sort into higher-quality institutions, estimation by OLS may overstate the effects of the institutions and produce an upward-biased estimate of β_3 (Angrist & Krueger, 2001). If preschool quality is endogenous, we expect the estimated β_3 to be smaller when applying the IV method than when we estimate by OLS. Given the concern about endogenous preschool quality, we do, in addition to OLS analyses, sensitivity analyses using the IV method.

We introduce instrumental variables that are highly correlated with the quality of the preschool institution but unrelated to child outcomes. The IVs allow us to study the Local Average Treatment Effects for those individuals whose treatment status is influenced by changes in the exogenous regressors (see Imbens & Angrist, 1994). The eight regressors that we investigated as regressors are: (1)-(5) overall characteristics of preschools at a larger geographical area for the five quality measures; (6) information about the political majority of the local government before preschool enrolment; (7) demographic changes in the preschool population across municipalities over the preceding years; and (8) whether the municipality has a GAPS policy at the time of preschool enrolment. We assess the strength of these eight instruments in first-stage t-statistics. For validity, we cannot directly

test the exclusion restriction (Angrist et al., 1996); instead we discuss the validity in each of the cases and perform tests of over-identifying restrictions.

Our first five instruments for the preschool institution's quality use the overall characteristics of preschools in a larger geographical area. In our preferred specification we use the county level. This instrumental variable approach exploits the variation in preschool institutions across 14 counties. For this procedure of using aggregates of our explanatory variables in larger geographical entities as an instrument for the explanatory variables at the local level, we follow Dustmann & Preston (2001), Heckman, Layne-Farra & Todd (1995) and Card & Krueger (1996). The validity of the instruments relies on the presumption that although preschool quality in the larger geographical area does not directly affect individual child outcomes, it is nevertheless related to preschool quality in the institution that the child attends. Parents' influence on the *overall* preschool quality in the county is marginal, and it is unlikely that parents move across counties because of preschool characteristics.⁵¹ Parental sorting *within* the county does not affect the overall composition in the county and thus does not affect validity. To ensure the strength of the instrument, we need not only a high correlation between county-level quality and specific institution-level quality but also variation in the preschool institutions across counties. The county aggregates of the five indicators show variation across Danish counties (see Table IV-8). However, for identification of possible quality effects, it is also important that there is some variation within the municipality.

We base the sixth instrument on the percentage of conservative votes for the local government prior to preschool enrolment.⁵² Given a stronger valuation of the family as childcare provider, a more conservative government may provide fewer resources and potentially lower quality for preschooling. Quality variations across municipal majorities appear in Table IV-9.⁵³ We do not expect that parents move to another municipality due to changes in the local political majority.

⁵¹ The mobility across county borders in Denmark is low; only around 3% of individuals in the labour force move across county borders per year (Deding & Filges, 2004). Simonsen (2005) shows a limited movement to and from municipalities providing generous childcare policies. Thus we do not worry about between-area sorting resulting from parents moving to other counties for better daycare.

⁵² A Conservative vote is defined as voting for the Liberal Democratic Party or the Progress Party. Data on the 1993 municipal elections are provided by Statistics Denmark (2006b).

⁵³ In conservative municipalities a larger share of children are born to native Danish parents, live in wealthier and better educated households and in larger (nuclear) families.

The seventh instrument, demographic changes in the population of preschool age children (Table IV-10), comes from municipal changes in the size of the population of 3-year-old children between 1995 and 1996 at the beginning of the studied preschool period (Statistics Denmark, 2006a). For example, a sudden positive change in the population size may affect the quality of preschool offerings, e.g. if preschools have to take care of more children, or if the waiting lists increase. Parents cannot anticipate a shock to the population size, and given the costs of moving to a new municipality, we expect that few of them move as a result (see also Datta Gupta & Simonsen, 2010, p. 33). As municipalities have more time to plan for this, unpredicted crowding will be less severe in primary and secondary schooling.

Finally, the eighth instrument is the presence or absence of a municipality policy of guaranteeing access to preschool (GAPS). We find that quality indicators vary somewhat with respect to the availability of GAPS in the previous year (Table IV-11), suggesting that GAPS is a valid instrument for quality in the subsequent year. We do not find any reverse causality between quality indicators in previous years and the provision of GAPS. Datta Gupta & Simonsen (2010), who use this instrument for the choice of preschool versus family care, show that there is no evidence that parental location choice is biased towards municipalities providing GAPS. A higher preschool take-up rate is expected in municipalities with GAPS. When a municipality provides GAPS, it is obliged to provide the necessary number of slots, thereby lengthening the waiting list if there is an increase in demand. This finding is in line with Datta Gupta & Simonsen (2010), who argue that if sufficient funds are available and municipalities try to reduce excess capacity, we could expect (for a given year) most of the variation in the provision of GAPS to stem from unexpected variations in demand, such as variations in the cohort size.

6. Estimation results

Our main results using OLS appear in section 6.1. And the following section, 6.2, contains results from the Two-Stage Least Square analysis (2SLS) using instrumental variables. In section 6.3, we present a discussion of the size of effects.

6.1. OLS estimates

Table IV-3 presents results for ten OLS estimation models. In model 1, to test whether our five indicators catch separate variation in the outcome, we include them jointly. Then we study the indicators separately (models 2, 4, 6, 7 and 9), presenting only output of those estimates that are significant. Finally, for each indicator, we test whether the effects vary across subgroups that are at risk of school failure (models 3, 5, 8 and 10).

A set of control variables for individual child and family characteristics is included in all models: the estimated coefficients show the expected signs, are almost all significant, and vary little across model specifications. The estimated models explain about 14.7% of the variation in Danish test outcomes, which is a remarkably high share in light of the long time between the preschool and testing periods. Tests for multicollinearity show that the included factors cover distinct parts of the outcome variation – cross-correlation is small, while calculations of the *variance inflation factors* (VIF) are all far below the threshold value of 10. Sudden shocks in the intermediary period between preschooling and testing at the end of secondary school are of no significant importance, as an assessment of the occurrence of 6% divorces in the studied sample shows.

The length of a preschool period is not significant. Our study of the effect of preschool quality on child outcomes is confined to children who attended a preschool. Children who did not attend a preschool would be likely to benefit from attending one, particularly one of a higher quality. That we cannot include this group of children introduces a possible downward bias of our estimates.

We initially study the effects of the five indicators jointly, which may be due to the limited variation in Denmark's near-universal daycare system. Three of five indicators – the staff-to-child ratio, the share of male teachers, and the share of trained teachers – show a significant positive relationship with the child outcome. The share of ethnic minority staff and staff stability are not significant. As several of the effects are significant in the joint estimation, our indicators cover distinct quality aspects. Separate tests for the indicators (models 2, 4, 6, 7 and 9) reconfirm the importance of preschool quality.

[For Table IV-3, please turn the page.]

Table IV-3: OLS regression output including all control variables

Dependent variable: Written final exam in Danish language skills at 9 th grade (end of primary schooling) in 2008, standardized scores (mean 100, std. dev. 15)										
Estimated Model	1	2	3	4	5	6	7	8	9	10
Preschool quality										
Number of staff members per child	2.55* [1.72]	2.06 [1.44]	-0.61 [-0.33]							
Share of male staff members	1.89** [2.16]			1.76** [2.02]	0.02 [0.02]					
Share of trained staff	1.93** [2.35]					1.59** [2.00]				
Share of ethnic minority staff	1.72 [1.30]						1.66 [1.25]	2.61* [1.87]		
Stability of staff	0.29 [0.39]								-0.06 [-0.09]	-0.35 [-0.50]
Subgroup interaction effects										
Ethnic minority children X quality indicator								-7.50** [-2.13]		4.81* [1.92]
Boys X quality indicator										

(continued on next page)

Table IV-3: OLS regression output including all control variables (*continued*)

Estimation Model (<i>continued</i>)		1	2	3	4	5	6	7	8	9	10
Peer group composition		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Child characteristics											
Length of attendance at last preschool (<i>in years</i>)		0.75	0.80	0.79	0.75	0.75	0.69	0.78	0.80	0.76	0.77
		[1.47]	[1.56]	[1.54]	[1.47]	[1.47]	[1.36]	[1.53]	[1.57]	[1.48]	[1.50]
Length of attendance at last preschool, squared term		-0.12	-0.13	-0.13	-0.12	-0.12	-0.11	-0.13	-0.13	-0.12	-0.12
		[-1.20]	[-1.32]	[-1.30]	[-1.18]	[-1.19]	[-1.07]	[-1.28]	[-1.33]	[-1.21]	[-1.24]
Gender (<i>boy = 1</i>)		-6.19***	-6.19***	-7.44***	-6.19***	-6.61***	-6.20***	-6.20***	-6.20***	-6.20***	-6.19***
		[-34.68]	[-34.64]	[-12.45]	[-34.69]	[-26.89]	[-34.70]	[-34.70]	[-34.71]	[-34.70]	[-34.69]
Age (<i>in 1998</i>)		-2.19***	-2.17***	-2.17***	-2.18***	-2.18***	-2.19***	-2.17***	-2.16***	-2.17***	-2.17***
		[-9.00]	[-8.90]	[-8.93]	[-8.95]	[-8.95]	[-8.99]	[-8.88]	[-8.86]	[-8.91]	[-8.90]
Ethnic minority (<i>Yes = 1</i>)		-3.30***	-3.31***	-3.32	-3.30***	-3.30***	-3.29***	-3.31***	-2.76***	-3.31***	-6.47***
		[-7.90]	[-7.94]	[-7.95]	[-7.91]	[-7.91]	[-7.88]	[-7.93]	[-5.65]	[-7.92]	[-3.86]
Family characteristics		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant		109.79***	111.08***	111.75***	111.46***	111.67***	111.12***	111.57***	111.41***	111.70***	111.85***
		[47.94]	[49.53]	[49.46]	[50.05]	[50.31]	[49.77]	[50.14]	[50.05]	[49.35]	[49.32]
Observations		30,444	30,444	30,444	30,444	30,444	30,444	30,444	30,444	30,444	30,444
Adjusted R ²		0.147	0.146	0.147	0.146	0.147	0.147	0.146	0.147	0.146	0.146

Note: Baseline category for Highest parental education = secondary education / vocational training; Robust t-statistics in parentheses, clustered at individual elementary school level (=1,421 clusters); * significant at 10% level ($|t|>1.64$), ** at 5% level ($|t|>1.96$); *** at 1% level ($|t|>2.58$). For the full OLS regression output, please refer to

Next we focus on the interaction between our quality indicators and a set of dummies for disadvantaged subgroups.⁵⁴ Subgroups of boys and ethnic minority children show significant interactions with our indicators. We find that a higher staff-to-child ratio is particularly positive for boys (model 3). Furthermore, as expected, we find that the presence of more male staff has a positive effect on outcomes for boys (model 5). Indeed, both subgroup effects drive the overall positive effects of more staff per child and a higher share of male staff, respectively, but the main effects are not significant when we include the interaction terms. Moreover, although children benefit from more trained staff, no significant variation appears across subgroups (model 7).

In addition, while the overall mean effect of staff stability has no significant long-term effect on Danish language skills, ethnic minority children benefit from a more stable staff (model 10): This effect is significant at a 10% level. The importance of staff stability for ethnic minority children could support the ‘attachment theory’ (Bowlby, 1969). This theory predicts that stability among the preschool staff might partially compensate for a lack of stability in the environment of ethnic minority children due to, for example, a recent migration experience or insufficient integration of the family into the local setting.

Ethnic minority children benefit significantly less from a higher share of ethnic minority staff (model 8). The negative interaction term between being an ethnic minority child and the share of ethnic minority staff is somewhat surprising. Initially, one might expect that ethnic minority staff would be better able to understand the needs of ethnic minority children. However, our data shows that ethnic minority staff has less work experience, has lower hourly wages, is less likely to have formal teacher training, and has less general education. Another explanation might be that preschools with more ethnic minority staff attract more ethnic minority children who are from disadvantaged families. Correlations between the five peer group indicators and the share of ethnic minority staff confirm an exceptionally strong correlation with the share of ethnic minority children in the preschool (about 40%).⁵⁵

⁵⁴ Table 3 shows only estimates of significant subgroup interaction effects plus the main effects. Regression outputs for all subgroup quality interactions, including those that did not show any significant effects, are available upon request.

⁵⁵ The negative coefficients of the interaction term could also be a result of reverse causality, especially if disadvantaged children are concentrated in preschools with higher quality, e.g. with more ethnic minority staff being assigned to disadvantaged groups. Indeed, we find that ethnic minority children are concentrated in centres with more ethnic minority staff.

We investigated whether boys from the subgroups of low-income families, ethnic minority families or single-parent households are affected differently by preschool quality than are other groups.⁵⁶ We find that ethnic minority boys benefit more from higher levels of all quality indicators than ethnic Danish boys or ethnic minority girls, except for higher shares of ethnic minority staff. Moreover, poor boys benefit more from a higher share of male staff than poor girls.

Estimating the same models using a set of other child outcomes confirms those effects. Long-term effects of high-quality preschooling on *oral* Danish skills are closely in line with those on *written* Danish skills. The presumption that high-quality preschooling is most influential for language development is confirmed also by written English and mathematics problem-solving outcomes - effects on English skills being largely in line with those on Danish skills, while cognitive development appears to be not affected. English outcomes are particularly positively related to having more non-native staff in a preschool.

It could be that estimating the OLS regression at the means may be sensitive to outliers or contribute to even out opposite properties at the end of the tails in the distribution of child outcomes (see Koenker & Basset, 1978). We estimate quantile regressions of the full model (1) over different deciles to test for the distribution of effects.

[For Table IV-4, please turn the page.]

⁵⁶ Output with interactions for disadvantaged boys is available upon request.

Table IV-4: Simultaneous quantile regression

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Number of staff members per child	2.69* [1.83]	1.37 [1.54]	0.00** [2.20]	1.09 [0.96]	0.00 [-0.24]	0.00 [-1.12]	0.00 [0.04]	1.4 [1.36]	0.00 [-0.56]	0.00 [-0.00]
Share of male staff members	1.03 [1.41]	1.42 [1.29]	0.00*** [-3.93]	2.54*** [3.44]	0.00*** [-3.91]	0.00 [-1.59]	0.00 [0.07]	0.43 [0.59]	0.00 [-0.69]	0.00 [-0.00]
Share of trained staff	0.87 [1.40]	0.96 [1.62]	0.00 [0.56]	1.99*** [3.03]	0.00*** [-3.01]	0.00 [-1.50]	0.00* [1.90]	1.01 [1.04]	0.00 [0.04]	0.00 [0.00]
Share of ethnic minority staff	1.33 [1.00]	0.48 [0.48]	0.00 [0.40]	1.65 [1.09]	0.00 [0.54]	0.00 [-0.08]	0.00 [-0.23]	0.49 [0.58]	0.00 [-0.19]	0.00 [-0.00]
Stability of staff	-0.02 [-0.05]	0.00 [-0.00]	0.00 [0.97]	-0.04 [-0.08]	0.00 [1.28]	0.00 [0.37]	0.00 [-0.39]	0.58 [1.06]	0.00 [-0.63]	0.00 [-0.00]

Note: Effect estimates for preschool quality aspects when running the OLS model (1) by decile, including peer composition, child characteristics and family background factors. Test scores are standardized to mean 100, standard deviation 15; level t-statistics based on bootstrapped errors in brackets; * significant at 10% level, ** at 5% level, and *** at 1%.

Quantile regressions reconfirm that three of the five quality indicators are significantly predicting child outcomes. A higher number of staff members per child is important at the lower end of the distribution, whereas the shares of male and trained staff members are more important for children with outcomes just below the median.

6.2. IV estimates

If the allocation of preschool slots through waiting lists has a partially random character, we control for various background factors that account for parental choices, preschool compositions and children's initial ability: these OLS results can be interpreted as causal. However, if we have not fully accounted for the treatment allocation process, OLS may produce a biased estimate of the true effect. If high-ability parents (with high-ability children) sort into high-quality preschools, the OLS estimate overestimates the true effect. Consequently, OLS estimates may be biased upward.

We now turn to the IV analyses to do a sensitivity check on whether effects may indeed be biased upward. We use the following eight instruments: five variables representing county-level aggregates of our five quality indicators, a variable for the share of conservatives and liberals on the municipal council, a dummy for whether the municipality provided GAPS for 3-6-year-olds at the time of enrolment between 1995 and 1998, and a variable for the relative change in the population of 3-year-old children in 1995. We find significant relationships between the indicators and all eight instrumental variables (see first-stage results in Table IV-5.a). For most specifications, the IVs are jointly significant and provide F-statistics above the recommended level of 10 (Baum et al., 2003; Staiger & Stock, 1997). We use the set of all eight instruments on each of the five possibly endogenous indicators. Table IV-5.b compares output for OLS and 2SLS regression outputs for the estimation model, which includes all five indicators jointly (IV model 1) and five separate models for each of the indicators (IV models 2-6).⁵⁷

For each of the 2SLS models we perform post-estimation tests to check whether 2SLS provides a better fit to the data than an OLS specification. First, we use a Wooldridge test (1995) to examine whether the quality indicators for which we want to use instrumental variables are exogenously determined.⁵⁸ Second, we use a

⁵⁷ IV regressions using subsets of instruments are largely consistent with regressions using the full set of IVs.

⁵⁸ In a separate sensitivity study we test and confirm that the full set of IVs is jointly strong enough to predict each of the potentially endogenous regressors, following a procedure as described by Angrist & Pischke (2008, p. 218).

Hausman test (1978) to check for systematic differences in the estimated coefficients in each IV and OLS specification (see Table IV-5.c).

For the specification of the full model (1), the results of the 2SLS estimation show that the quality indicators are not significant. According to Wooldridge's test of exogeneity, the instrumented variables are exogenous across all specifications. Moreover, Hausman tests do not point to systematic differences in the OLS and 2SLS estimates, suggesting that OLS is the preferred estimation method. The same holds when we analyse our five quality indicators separately. However, lack of power of the instruments could explain the cases in which a Hausman test detects no differences between the 2SLS and OLS coefficients.

[For Table IV-5, see following page.]

Table IV-5: 2SLS regression output, applying all 8 instrumental variables

a 1st stage results

Dependent variable: Written final exam in Danish language skills at 9 th grade (end of primary schooling) in 2008, standardized scores (mean 100, std. dev. 15)										
Estimate d Model	1	2	3	4	5	6	7	8	9	10
Preschool quality	Staff-to- child ratio	Share of male staff members	Share of trained staff	Share of ethnic minority staff	Stability of staff	Staff-to- child ratio	Share of male staff members	Share of trained staff	Share of ethnic minority staff	Stability of staff
Instrumental variables										
County aggregate: Staff-to- child ratio	0.77*** [12.21]	0.02 [0.21]	0.21** [1.68]	0.09* [1.75]	0.74*** [5.01]	0.77*** [12.21]	0.02 [0.21]	0.21* [1.68]	0.09* [1.75]	0.74*** [5.01]
County aggregate: Share of male staff	0.18*** [3.76]	0.89*** [11.00]	-0.22*** [-2.33]	-0.06 [-1.30]	-0.52*** [-4.46]	0.18*** [3.76]	0.89*** [11.00]	-0.22** [-2.33]	-0.06 [-1.30]	-0.52*** [-4.46]
County aggregate: Share of trained staff	0.02 [1.34]	0.10*** [3.55]	0.99*** [27.13]	0.04* [1.73]	-0.03 [-0.77]	0.02 [1.34]	0.10*** [3.55]	1.00*** [27.13]	0.04* [1.73]	-0.03 [-0.77]
County aggregate: Share of ethnic minority staff	-0.15*** [-4.78]	-0.02 [-0.32]	0.10** [1.73]	0.58*** [15.47]	0.07 [1.08]	-0.15*** [-4.78]	-0.02 [-0.32]	0.10* [1.73]	0.58*** [15.47]	0.07 [1.08]
County aggregate: Stability of staff	0.23*** [5.55]	0.05 [0.67]	-0.12 [-1.47]	0.10*** [2.88]	0.54*** [5.25]	0.23*** [5.55]	0.05 [0.67]	-0.12 [-1.47]	0.10*** [2.88]	0.54*** [5.25]
Municipal election 1993	-0.01*** [-7.34]	0.00 [0.61]	-0.00** [-2.49]	-0.01*** [-6.57]	0.00* [1.88]	-0.01*** [-7.34]	0.00 [0.61]	0.00** [2.49]	-0.00*** [-6.57]	0.00* [1.88]
GAPS 1995 (for age 3-6)	-0.01*** [-7.43]	-0.00** [-2.61]	-0.01*** [-3.75]	0.00 [-1.60]	0.02*** [10.38]	-0.01*** [-7.43]	-0.00*** [-2.61]	-0.01*** [-3.75]	0.00 [-1.60]	0.02*** [10.38]
Change child pop. 1995-96 (age 3)	0.00*** [-6.68]	-0.00*** [-8.66]	0.00*** [4.48]	-0.00** [-2.13]	0.00*** [5.91]	0.00*** [-6.68]	-0.00*** [-8.66]	0.00*** [4.48]	-0.00** [-2.13]	0.00*** [5.91]

(continued on next page)

a 1st stage results (continued)

Estimate d Model (continued)	1	2	3	4	5	6	7	8	9	10
F-test for joint significance of the instruments	F=20.65	F=35.23	F=24.59	F=18.08	F=15.48	F=20.65	F=35.23	F=24.59	F=18.07	F=15.48
	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00
	Strong IVs					Strong IVs	Strong IVs	Strong IVs	Strong IVs	Strong IVs
Over-identifying restrictions for 2nd stage regression: Sargan (1958)						$\chi^2=12.40$	$\chi^2=9.53$	$\chi^2=15.72$	$\chi^2=9.63$	$\chi^2=11.21$
						p=0.09	p=0.22	p=0.03	p=0.21	p=0.13
						IVs valid	IVs valid	IVs not valid	IVs valid	IVs valid

Note: Robust z-statistics in parentheses; * significant at 10% level ($|t|>1.64$), ** at 5% level ($|t|>1.96$); *** at 1% level ($|t|>2.58$). A critical F value for the test on joint significance of the IVs is 10 (as stated, for example, in Baum et al., 2003). The test for the over-identification is based on an estimation without robust standard errors and in which weights are forced; it checks the coherency of the chosen set of instruments rather than their validity (cf. Parente & Santos Silva, 2012)

b 2nd stage results

Dependent variable:

Written final exam in Danish language skills at 9th grade (end of primary schooling) in 2008, standardized scores (mean 100, std. dev. 15)

	1		2		3		4		5		6	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Preschool quality												
Staff-to-child ratio	2.55*	-5.8	2.06	11.55								
	[1.72]	[-0.35]	[1.44]	[1.25]								
Share of male staff	1.89**	5.37			1.76**	9.28						
	[2.16]	[0.41]			[2.02]	[1.63]						
Share of trained staff	1.93**	1.28					1.59**	-0.19				
	[2.35]	[0.22]					[2.00]	[-0.04]				
Share of ethnic minority staff	1.72	29.88*					1.66	21.61				
	[1.30]	[1.72]					[1.25]	[1.63]				
Stability of staff	0.29	-8.64									-0.06	-8.73
	[0.39]	[-0.79]									[-0.09]	[-1.39]
Peer group	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Family characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	109.79***	119.06***	111.08***	111.26***	111.46***	113.21***	111.12***	114.49***	111.57***	113.3	111.70***	120.24***
	[48.94]	[13.46]	[49.53]	[30.88]	[50.05]	[41.01]	[49.77]	[34.78]	[50.14]	[41.11]	[49.35]	[24.56]
Observations	30,444	20,501	30,444	20,501	30,444	20,501	30,444	20,501	30,444	20,501	30,444	20,501

Note: Robust t-statistics (z-statistics for IV) in parentheses, clustered at individual elementary school level (=1,421 clusters); * significant at 10% level ($|t|>1.64$), ** at 5% level ($|t|>1.96$), *** at 1% level ($|t|>2.58$).

c Post-estimation model tests

Estimated IV Model	1	2	3	4	5	6
Wooldridge's test of exogeneity of instrumented variables	F=1.03 / p=0.40	F=1.18 / p=0.28	F=1.55 / p=0.21	F=0.08 / p=0.78	F=2.30 / p=0.13	F=2.04 / p=0.15
	Exogenous	Exogenous	Exogenous	Exogenous	Exogenous	Exogenous
Hausman test for no systematic differences between IV and OLS	$\chi^2=35.89$ / p=0.06	$\chi^2=29.17$ / p=0.08	$\chi^2=31.17$ / p=0.05	$\chi^2=27.60$ / p=0.12	$\chi^2=31.21$ / p=0.05	$\chi^2=30.69$ / p=0.06
	Marginally significant difference	No significant difference	Significant difference	No difference	Significant difference	Marginally significant difference
Interpretation of estimation results	Prefer OLS	Prefer OLS	Prefer OLS	Prefer OLS	Prefer OLS	Prefer OLS

6.3. Size of effects

To compare the effects across indicators, we study the effects from the OLS and IV estimations.⁵⁹ The IV estimates are somewhat higher than the OLS estimates. We report the size of our estimated effects as intervals given by the OLS and IV estimates, respectively. The main effects for a unit change in the five quality indicators range around 2-4 standardized points in the OLS estimations, and somewhat larger, about 0-19 standardized points, in the IV estimations. For example, an increase in the staff-to-child ratio of around one standard deviation (0.065, or a 30% increase in teacher resources) implies an increase in the standardized score of 0.17-0.72 (corresponding to an interval between $2.55 \cdot 0.065$ and $11.1 \cdot 0.065$). For the average pupil with an average test score at around 6.6 on the Danish grade scale, this finding translates into a grade increase of 0.01-0.05. Thus a substantial increase in teacher resources is associated with a rather modest increase in grades. Nonetheless, one should note that almost 40% of the pupils received a grade of 7 in written Danish. For comparison, “raw” average grade points (not controlling for background factors) often vary by about 0.1-0.5 across schools in the same municipality. For boys, the corresponding increase in grades doubles in the OLS specification (0.02 compared to 0.01), while interaction terms are insignificant in the IV specification.

An increase in the male share of the staff of one standard deviation (from 10% to 22%) implies an increase by 0.01-0.06 grade points (corresponding to OLS and IV estimates, respectively). Again, the effect is higher for boys in the OLS specification. Thus the effect doubles for boys. Moreover, for poor boys, the higher share of male staff is associated with a difference in grade points of around six times the overall effect found in the OLS estimations.

An increase in trained staff by one standard deviation of 0.12 (from 45% to 57%) implies an increase in the standardized score of 0.23 ($=1.9 \cdot 0.12$), corresponding to 0.02 grade points.

We also found significant effects of the share of ethnic minority teachers in the staff. An increase in the share of ethnic minority teachers from 4% to 7% is associated with a 0.02-0.09 higher grade point in written Danish.

⁵⁹ Each coefficient of the explanatory variables is recalculated in terms of how much a standard error change in the variable changes the outcome variable. The effects of each explanatory variable can therefore be compared directly. The standardized beta coefficients are based on the same OLS regression specification. As the beta option cannot be combined with clustering, we instead use the robust option for the standard errors. The full outputs with standardized beta coefficients are available upon request.

We found no significant effect of higher staff stability in general. However, for ethnic minority children, we find a substantial effect of more stability: an increase in staff stability of one standard deviation (reduction in teacher turnover per year from 33% to 20%) translates into a 0.05 higher grade on the Danish grade-point scale.

For comparison, being a girl is associated with a 0.42 higher grade in written Danish, having a parent with a middle or higher education is associated with a 0.35 higher grade, and coming from a two-parent family results in a 0.10 higher grade for the average child.

7. Discussion and conclusion

Policymakers in many OECD countries are considering extending early childhood care and education towards systems of (near) universal access through substantial public subsidies. More universal access to early investments can potentially increase a country's human capital and reduce inequality in later child outcomes. This chapter adds unique exploratory evidence for quality factors in preschools and finds evidence of peer group effects stemming from the preschool's share of children with less educated parents.

Preschool quality comprises a wide range of factors. We investigate five indicators of quality: (1) the staff-per-child ratio, (2) the share of male staff members, (3) the share of trained staff, (4) the share of ethnic minority staff members, and (5) the stability of staff per institution compared to the previous year. While these indicators do not capture all aspects of quality in preschools, the value of our indicators lies in their being objectively measurable and comparable across preschools and municipalities. Testing a set of instrumental variables to identify potential estimate biases does not produce more consistent estimates than OLS regressions.

Overall, we find significant relationships between our preschool quality indicators and children's Danish language skills, even after inclusion of various controls for the child's background. Even 10 years after children's preschool experience, the effects are still visible, thus demonstrating the importance of quality. The number of staff per child and the shares of male staff, trained staff and ethnic minority staff all have a significant, positive impact on child development. Boys benefit more from a higher number of staff per child and a higher share of male staff, whereas ethnic minority children benefit from higher staff stability. These findings indicate

that preschool quality helps equalize outcomes for boys and ethnic minority children.⁶⁰

Our evidence on quality impacts results from an evaluation of experienced quality variation across children who attended preschools in the second half of the 1990s. Thus the results do not necessarily apply to the 10-15% of children in 1998 who attended educational formats other than preschool or age-integrated institutions. Children with alternative childcare experiences would most likely benefit from high-quality preschooling. In that case, we may underestimate the true effects. Moreover, as effect evaluation methods rule out controlling for children's development after preschool, we cannot identify which part of child outcomes is the result of preschool quality and which part is the result of their subsequent schooling.

Our results are in line with the literature that finds positive long-term impacts of preschool attendance (e.g. Baker et al., 2008; Berlinski, Galiani, & Gertler, 2009; Berlinski, Galiani, & Manacorda, 2008; Deming, 2009; Dumas & Lefranc, 2010; Felfe & Lalive, 2010; Havnes & Mogstad, 2011). While this body of research evaluates the effects of preschooling in general, we add new evidence of the effects of variation in preschool quality. Moreover, we find that the effects are not the same for all children across heterogeneous socioeconomic backgrounds and characteristics. For example, our results provide some support for the earlier finding that boys benefit more from higher quality of childcare (Datta Gupta & Simonsen, 2010). In contrast to Chetty et al. (2011), who argue that the effect of higher quality on cognitive development fades out after a few years, we find persistent quality returns even at the end of primary school. Our study shows that, even after 10 years of later schooling, the quality of daycare has a *lasting* – albeit numerically modest – impact on children's linguistic outcomes.

⁶⁰ Our results are based on linear estimations. A non-linear estimation of subgroup effects might reveal more information about policy targeting to specific disadvantaged subgroups.

Appendix IV

Table IV-6: Descriptive statistics of control variables

Control variable	N	Mean	Std. dev.	Min.	Max.
Share of boys in attended preschool	30,444	0.519	0.067	0.2	0.9
Share of ethnic minority children	30,444	0.058	0.103	0.0	1.0
Share of children with single parents	30,444	0.204	0.107	0.0	0.8
Share of children with less educated parents	30,444	0.122	0.090	0.0	0.6
Share of children from poor households	30,444	0.043	0.057	0.0	0.6
Length of attendance at last preschool (in years)	30,444	2.661	0.862	1.0	4.0
Gender (boy = 1)	30,444	0.499	0.500	0.0	1.0
Age (in 1998)	30,444	6.076	0.342	4.0	9.0
Ethnic minority (no parent a Danish-born citizen)	30,444	0.062	0.241	0.0	1.0
Parenthood (single parents = 1)	30,444	0.196	0.397	0.0	1.0
Number of children in household (in 1998)	30,444	2.193	0.759	0.0	10.0
Mother's age (in 1998)	30,444	34.968	4.569	21.0	53.0
Maternal unemployment (annual days at 8 hrs. in 1998)	30,444	9.019	22.176	0.0	125.0
Highest parental education (elementary)	30,444	0.101	0.301	0.0	1.0
Highest parental education (secondary/voc. training)	30,444	0.483	0.500	0.0	1.0
Highest parental education (low tertiary)	30,444	0.081	0.274	0.0	1.0
Highest parental education (mid. tertiary/bachelor)	30,444	0.221	0.415	0.0	1.0
Highest parental education (high tertiary)	30,444	0.114	0.318	0.0	1.0
Gross hh-income (logarithmic scale, in 1998)	30,444	13.822	1.064	3.5	19.3

Table IV-7: Danish grades and the averages of their standardized scores

Danish exam grade		Average standardized scores	Frequency
-3	Unacceptable	50.15	24
0	Inadequate	65.86	279
2	Adequate	76.33	2,326
4	Fair	86.80	7,834
7	Good	102.50	11,746
10	Very good	118.21	6,707
12	Excellent	128.68	1,528

Table IV-8: Aggregated preschool quality averages across Danish counties

County	Municipal. ^a	Obs.	Staff- per- child ratio	Share of male staff (0-1 scale)	Share of trained staff (0-1 scale)	Share of ethnic minority staff (0-1 scale)	Stability of staff
	N	N	mean	mean	mean	mean	mean
Bornholm	5	332	0.249	0.106	0.365	0.023	0.697
Copenhagen Capital Region ^b	20	6,368	0.250	0.155	0.421	0.074	0.650
Frederiksborg	19	2,672	0.230	0.109	0.453	0.047	0.692
Fyn	32	2,096	0.230	0.127	0.472	0.021	0.662
North Jutland	27	1,933	0.203	0.098	0.490	0.025	0.653
Ribe	14	820	0.200	0.063	0.505	0.018	0.715
Ringkjøbing	18	824	0.190	0.102	0.440	0.013	0.661
Roskilde	11	1,459	0.231	0.075	0.376	0.052	0.700
South Jutland	23	1,923	0.214	0.078	0.478	0.062	0.705
Storstrøm	24	1,674	0.218	0.100	0.460	0.022	0.712
Vejle	16	2,207	0.206	0.123	0.479	0.025	0.661
Viborg	17	1,015	0.211	0.100	0.476	0.019	0.668
West Zealand	23	2,334	0.207	0.096	0.453	0.035	0.674
Århus	26	4,787	0.247	0.158	0.467	0.039	0.649
Range of mean values			0.190- 0.250	0.063- 0.158	0.365- 0.505	0.013- 0.074	0.649- 0.715

Note:

- a) Geographic division before 2006/2007, 275 municipalities in total; No information available for 7 rural/island municipalities with small child populations
- b) Includes the municipal regions Copenhagen and Frederiksberg

Table IV-9: Quality difference of conservative municipalities, 1993 election

	Observations	Staff-per-child ratio	Share of male staff (0-1 scale)	Share of trained staff (0-1 scale)	Share of ethnic minority staff (0-1 scale)	Stability of staff
	N	mean	mean	mean	mean	mean
Non-conservative majority	21,518	0.231	0.125	0.456	0.045	0.668
Conservative majority	8,926	0.220	0.111	0.447	0.036	0.673
5% t-test		difference significant	difference significant	difference significant	difference significant	difference significant

Table IV-10: Municipal child population aged 3, %-change 1995/96 by centile

	Observations	Staff-per-child ratio	Share of male staff (0-1 scale)	Share of trained staff (0-1 scale)	Share of ethnic minority staff (0-1 scale)	Stability of staff
	N	mean	mean	mean	mean	mean
1st centile	3,102	0.227	0.104	0.435	0.069	0.673
2nd centile	3,021	0.227	0.112	0.471	0.037	0.663
3rd centile	3,108	0.218	0.124	0.425	0.052	0.672
4th centile	3,038	0.226	0.106	0.463	0.044	0.687
5th centile	3,157	0.257	0.174	0.451	0.048	0.623
6th centile	3,977	0.239	0.160	0.442	0.058	0.658
7th centile	2,057	0.224	0.121	0.451	0.044	0.686
8th centile	2,927	0.223	0.115	0.451	0.048	0.667
9th centile	3,113	0.216	0.096	0.473	0.032	0.673
10th centile	2,942	0.216	0.084	0.472	0.025	0.706
Value ranges		0.215-0.257	0.085-0.174	0.425-0.473	0.025-0.069	0.623-0.706

Table IV-11: Availability of guaranteed preschool place scheme in municipality

Provision of municipal guaranteed preschool scheme (GAPS) for 3-6-olds in 1995	Observations	Staff-per-child ratio	Share of male staff (0-1 scale)	Share of trained staff (0-1 scale)	Share of ethnic minority staff (0-1 scale)	Stability of staff
	N	mean	mean	mean	mean	mean
NO	5,817	0.243	0.129	0.448	0.048	0.647
YES	14,684	0.231	0.125	0.444	0.044	0.667
5% t-test		difference significant	difference significant	difference significant	difference significant	difference significant

Note: We use GAPS information provided by an inventory of Glavind (2007). Our indicator checks whether GAPS was provided to the preschool age group 3-6 in 1995, when the majority of the studied children enrolled at preschool. For a number of municipalities no information on GAPS was available, reducing the sample size to N=20499. There is no indication that the lack of GAPS information is systematically correlated with the treatment or outcome variables.

Table IV-12: OLS regression output including all control variables (*full output*)

Dependent variable: Written final exam in Danish language skills at 9 th grade (end of primary schooling) in 2008, standardized scores (mean 100, std. dev. 15)										
Estimated Model	1	2	3	4	5	6	7	8	9	10
Preschool quality										
Number of staff members per child	2.55* [1.72]	2.06 [1.44]	-0.61 [-0.33]							
Share of male staff members	1.89** [2.16]			1.76** [2.02]	0.02 [0.02]					
Share of trained staff	1.93** [2.35]					1.59** [2.00]				
Share of ethnic minority staff	1.72 [1.30]						1.66 [1.25]	2.61* [1.87]		
Stability of staff	0.29 [0.39]								-0.06 [-0.09]	-0.35 [-0.50]
Subgroup interaction effects										
Ethnic minority children X quality indicator								-7.50** [-2.13]		4.81* [1.92]
Boys X quality indicator										

(continued on next page)

Table IV-12: OLS regression output including all control variables (*full output*)
(*continued*)

Estimated Model (<i>continued</i>)										
Peer group composition										
Share of boys in attended preschool	0.9	1.09	1.06	1.08	1.1	1.03	1.05	1.08	1.11	1.12
	[0.68]	[0.82]	[0.80]	[0.82]	[0.83]	[0.78]	[0.80]	[0.82]	[0.84]	[0.85]
Share of ethnic minority children	-1.98	-1.45	-1.46	-1.51	-1.51	-1.18	-1.87	-1.28	-1.35	-1.23
	[-1.42]	[-1.06]	[-1.07]	[-1.11]	[-1.11]	[-0.87]	[-1.34]	[-0.90]	[-0.99]	[-0.92]
Share of children with single parents	-1.04	-0.84	-0.81	-1.08	-1.07	-0.74	-0.84	-1.04	-0.82	-0.88
	[-0.76]	[-0.62]	[-0.59]	[-0.79]	[-0.78]	[-0.54]	[-0.62]	[-0.76]	[-0.61]	[-0.65]
Share of children with less educated parents	-9.10***	-9.66***	-9.66***	-9.45***	-9.46***	-9.85***	-9.75***	-9.71***	-9.85***	-9.86***
	[-6.52]	[-6.93]	[-6.94]	[-6.77]	[-6.78]	[-7.09]	[-7.02]	[-6.99]	[-7.07]	[-7.08]
Share of children from poor households (<i>bottom quintile</i>)	4.22	4.66*	4.67*	4.76*	4.71*	4.85*	5.16*	5.13*	5.03*	5.13*
	[1.57]	[1.72]	[1.72]	[1.77]	[1.75]	[1.80]	[1.91]	[1.90]	[1.87]	[1.90]
Child characteristics										
Length of attendance at last preschool (<i>in years</i>)	0.75	0.8	0.79	0.75	0.75	0.69	0.78	0.8	0.76	0.77
	[1.47]	[1.56]	[1.54]	[1.47]	[1.47]	[1.36]	[1.53]	[1.57]	[1.48]	[1.50]
Length of attendance at last preschool, squared term	-0.12	-0.13	-0.13	-0.12	-0.12	-0.11	-0.13	-0.13	-0.12	-0.12
	[-1.20]	[-1.32]	[-1.30]	[-1.18]	[-1.19]	[-1.07]	[-1.28]	[-1.33]	[-1.21]	[-1.24]
Gender (<i>boy = 1</i>)	-6.19***	-6.19***	-7.44***	-6.19***	-6.61***	-6.20***	-6.20***	-6.20***	-6.20***	-6.19***
	[-34.68]	[-34.64]	[-12.45]	[-34.69]	[-26.89]	[-34.70]	[-34.70]	[-34.71]	[-34.70]	[-34.69]
Age (<i>in 1998</i>)	-2.19***	-2.17***	-2.17***	-2.18***	-2.18***	-2.19***	-2.17***	-2.16***	-2.17***	-2.17***
	[-9.00]	[-8.90]	[-8.93]	[-8.95]	[-8.95]	[-8.99]	[-8.88]	[-8.86]	[-8.91]	[-8.90]
Ethnic minority (<i>Yes = 1</i>)	-3.30***	-3.31***	-3.32***	-3.30***	-3.30***	-3.29***	-3.31***	-2.76***	-3.31***	-6.47***
	[-7.90]	[-7.94]	[-7.95]	[-7.91]	[-7.91]	[-7.88]	[-7.93]	[-5.65]	[-7.92]	[-3.86]

(*continued on next page*)

Table IV-12: OLS regression output including all control variables (*full output*)
(*continued*)

Estimated Model (<i>continued</i>)	1	2	3	4	5	6	7	8	9	10
Family characteristics										
Parenthood (<i>single parents = 1</i>)	-1.52*** [6.62]	-1.53*** [6.64]	-1.53*** [6.67]	-1.52*** [6.61]	-1.52*** [6.62]	-1.52*** [6.64]	-1.53*** [6.66]	-1.54*** [6.69]	-1.53*** [6.64]	-1.53*** [6.65]
Number of children in household (<i>in 1998</i>)	-1.15*** [10.89]	-1.16*** [10.94]	-1.16*** [10.94]	-1.15*** [10.94]	-1.15*** [10.94]	-1.16*** [11.01]	-1.16*** [10.98]	-1.15*** [10.88]	-1.16*** [10.99]	-1.16*** [10.99]
Mother's age (<i>in 1998</i>)	0.10*** [5.01]	0.10*** [5.08]	0.10*** [5.09]	0.10*** [5.03]	0.10*** [5.05]	0.10*** [5.07]	0.10*** [5.08]	0.10*** [5.07]	0.10*** [5.08]	0.10*** [5.06]
Maternal unemployment (<i>annual full-time days, 1998</i>)	-0.02*** [5.02]	-0.02*** [5.03]	-0.02*** [5.05]	-0.02*** [5.02]	-0.02*** [5.03]	-0.02*** [5.03]	-0.02*** [5.02]	-0.02*** [5.01]	-0.02*** [5.02]	-0.02*** [5.01]
Highest parental education (<i>elementary</i>)	-3.82*** [13.10]	-3.81*** [13.06]	-3.81*** [13.07]	-3.81*** [13.08]	-3.81*** [13.08]	-3.80*** [13.05]	-3.81*** [13.06]	-3.80*** [13.03]	-3.80*** [13.05]	-3.80*** [13.04]
Highest parental education (<i>low tertiary</i>)	2.79*** [9.26]	2.81*** [9.32]	2.81*** [9.33]	2.81*** [9.33]	2.82*** [9.35]	2.81*** [9.32]	2.81*** [9.34]	2.81*** [9.34]	2.82*** [9.35]	2.81*** [9.34]
Highest parental education (<i>mid. tertiary/bachelor</i>)	5.18*** [23.57]	5.22*** [23.75]	5.22*** [23.75]	5.22*** [23.72]	5.21*** [23.69]	5.22*** [23.69]	5.23*** [23.76]	5.23*** [23.75]	5.23*** [23.76]	5.23*** [23.77]
Highest parental education (<i>high tertiary</i>)	8.09*** [29.09]	8.14*** [29.37]	8.13*** [29.36]	8.13*** [29.30]	8.12*** [29.28]	8.17*** [29.48]	8.15*** [29.41]	8.14*** [29.38]	8.16*** [29.38]	8.16*** [29.38]
Gross hh-income (<i>logarithmic scale, in 1998</i>)	0.23*** [2.58]	0.21** [2.45]	0.21** [2.45]	0.22** [2.49]	0.22** [2.47]	0.21** [2.44]	0.21** [2.42]	0.21** [2.46]	0.21** [2.41]	0.21** [2.44]
Constant	109.79*** [47.94]	111.08*** [49.53]	111.75*** [49.46]	111.46*** [50.05]	111.67*** [50.31]	111.12*** [49.77]	111.57*** [50.14]	111.41*** [50.05]	111.70*** [49.35]	111.85*** [49.32]
Observations	30,444	30,444	30,444	30,444	30,444	30,444	30,444	30,444	30,444	30,444
Adjusted R ²	0.147	0.146	0.147	0.146	0.147	0.147	0.146	0.147	0.146	0.146

Note: Baseline category for Highest parental education = secondary education / vocational training; Robust t-statistics in parentheses, clustered at individual elementary school level (=1,421 clusters); * significant at 10% level ($|t|>1.64$), ** at 5% level ($|t|>1.96$); *** at 1% level ($|t|>2.58$).

Data Annotation

The dataset we used to study the long-term effects of Danish preschool quality was prepared with the financial support of the Centre for Strategic Research in Education and Competence Building. It is a unique, newly created administrative dataset - the first that linked preschool records with long-term child outcomes under the inclusion of preschool staff and parental records. I have carried out the full data preparation and run all analyses.

The syntax files, including several thousand syntax commands, have been programmed in a way to allow some flexibility to the choice of variables, years and records to be included in the final dataset, setting the ground for promising follow-up studies. A pre-selection has been made to keep the dataset manageable, given the use of very large population records. In all analyses, anonymity of individual records has been guaranteed.

For the initial preparation of the Danish data, I spent three months as a guest researcher at the Danish Institute for Local and Regional Government Research (KORA; former Danish Institute of Governmental Research, AKF). To prepare the dataset for analysis, I cleaned and merged more than ten separate registry datasets for the period 1995-2008 via AKF's access to Statistics Denmark servers, using a SAS software package, i.e. childcare registers with employment registers and childcare registers with basic registers to identify child and parental background characteristics, adding test records.

I then transferred the datasets to Stata, pre-formatted and merged them and standardized test scores for analysis. After leaving Denmark, I first used an AKF remote access to Statistics Denmark servers and later remote accesses to the Danish National Centre for Social Research and Department of Economics at the University of Copenhagen to run a number of estimation and sensitivity analyses.

Chapter V - Conclusions

Substantial private and public resources are invested in extending Early Childhood Care and Education provisions, with the general aim to facilitate maternal employment, e.g. through guaranteed daycare placement at the earliest possible ages. The quality of childcare is considered to be crucial for stimulating child development and, in particular, is expected to improve the chances of disadvantaged children to do better in their later schooling. The international literature offers ample theoretical support and promising indicative evidence from selected case studies on the potential benefits of high-quality childcare for child development and its equalizing effect on the distribution of chances in life.

Yet, the literature still tells rather little about what high-quality childcare entails and how resources should be invested to guarantee that the expected positive effects are maximised. With insufficient evidence about the impact about childcare quality aspects at hand, policymakers are likely to allocate public resources ineffectively; in particular, the chances of disadvantaged children may not be improved, so that inequalities of opportunity may persist.

This dissertation studies three major aspects of the quality of early childhood care and education investments: 1) the channel of delivery – home- versus centre-based provisions, 2) the kind of early childhood education approach – teacher-based or child-centred, and 3) the characteristics of the childcare staff – staff-child ratio, male staff ratio, trained staff ratio, ethnic minority staff ratio and staff turnover.

In this final chapter, I review the main findings of these studies, interpret their policy implications, point out their limitations and suggest next steps for future research. I then offer some closing remarks.

1. Main findings

The *first study* in this dissertation, chapter II, deals with estimating the effects on school readiness of attending centre-based childcare versus home-based provisions. I study those effects by using rich data from a national representative cohort study for the Netherlands and ordinary least squares regressions. To assess possible endogeneity biases, I apply province averages as instrumental variables (IV) and attempt a structural equation model (SEM). I find that child outcomes are more sensitive to the length and intensity of childcare spells in preschool kindergarten than daycare centres, indicating decreasing returns. However, my analysis does not support the positive expectations on the impact of centre- versus home-based childcare – the IV and SEM sensitivity analyses point into the same direction. Correspondingly, my estimates do not support worries about potential harms of centre-based childcare.

The *second study*, chapter III, studies how a child-centred early childhood education approach influences school readiness as compared to alternative childcare options. The rich dataset that is used for this study covers almost all children in the southern part of the Dutch province of Limburg. I apply ordinary least squares regressions and assess the validity of results by applying two alternative strategies: propensity score matching as well as instrumental variable estimations, which are based on early childhood education investment patterns of municipalities and management boards of preschool kindergartens. I find no evidence of the expected short-term effects on language or on cognitive development for children who attended a preschool with child-centred early childhood education as compared to children who attended preschools applying other early education programmes or none at all. In order to reach measurable benefits, the child-centred methods and their applications need to be intensified and extended to all disadvantaged groups of children. However, I find some evidence that children of less educated parents who are placed in a preschool with child-centred early childhood education tend to have higher language and cognitive outcomes.

The *third study*, chapter IV, assesses how children in Denmark vary in their development according to different childcare staff characteristics at preschool kindergartens in the late 1990s. Danish administrative childcare records have been merged for the first time with child outcomes from the final test of elementary schooling and a number of individual and family background characteristics. Childcare and employment records are used to generate five measures of preschool quality, all related to staffing decisions: the number of staff members per child; the shares of male staff, staff with a pedagogic education, and staff with an ethnic minority background and the stability of staff across years.

OLS analyses conditioning on a set of individual and parental characteristics show that four out of five indicators—a higher staff-to-child ratio, a higher share of male staff, a higher share of staff with formal preschool teacher training, and a higher share of ethnic minority teachers—are linked to significantly better performance in children's test results in Danish. Boys benefit more from preschool quality than girls. Ethnic minority children benefit from higher staff stability. To the degree that the final allocation of children is determined by waiting lists of municipalities, these OLS estimates can be assigned causal meaning. A sensitivity analysis, using an instrumental variable approach to account for possible endogeneity in the allocation mechanism, does not have strong enough results to confirm significant positive effects for any indicators of quality.

What remains is indicative evidence that supports the relevance of the length and intensity of childcare, the short-term impact of child-centred early childhood education for children of less educated parents, and the long-term effect of some staff-related quality aspects.

2. Interpretation of results

This empirical evidence on the effects of early childhood care and education investments still falls short of supporting the strong theoretical argument, for most of the studied quality characteristics and child outcomes. What may be the reason for this shortcoming? There are a number of possible factors: reasons regarding the assessed child outcome and the studied childcare quality characteristics as well as reasons regarding the applied empirical strategy.

A first consideration may be that the studied child outcomes are only proxies for the true child outcomes, reflecting only somewhat narrower domains of child development. The impact of the studied quality characteristics may, however, be reflected only partially in those narrower development domains. A lack of significant effects may thus be a result of not having been able to study the relevant development domains, because of a lack of data. Also, the child outcomes may proxy the true state of child development with a higher degree of measurement error. This problem reflects a general weakness of effect studies that rely on given test outcomes. The problem can be reduced by choosing child outcomes that cover different development domains, that are available on a larger scale, that have been repeatedly used for similar effect studies and that are regularly used and refined to assess child development. Within the data limitations, the strategies have been followed in this dissertation as far as possible.

Also, language outcomes that have been more of a focus of policy makers in recent years, e.g. with respect to integrating ethnic minority children and early education initiatives, may be less sensitive to preschooling than cognitive outcomes. Above that, it may take some time for effects to reveal themselves in measurable child outcomes. Following Cunha and Heckman (2007, 2008), who argue that early development gains may multiply the effects of later human resource investments, it may be that effects that still look weak at the beginning of the elementary school may facilitate the acquisition of later skills and this be measurable later on. In that regard we would expect stronger evidence for the third study, in which child outcomes at the end of elementary school are examined. However, whether effects become evident at a later stage depends on whether the quality of later education is in line with the quality of early childhood care and education so that effects fade out and no other distorting influences happen in the meantime, such as, for example, the child falling ill, losing a parent or other closely attached person, or parental divorce.

Another reason for the weak evidence may be that the studied children have been insufficiently exposed to the studied childcare quality characteristics (see, e.g. Leseman, 2002). This may be particularly the case if children have attended the respective childcare provision for only few sessions per week or short time periods; effects may only become visible with higher intensities or longer

experiences. This may be to a larger degree the case for the studied Dutch childcare provisions, which expose children to lower intensities and durations of childcare, than for the near-universal, more extensive Danish system. In particular, the study of the Dutch early childhood education approaches may be prone to such low exposure. Dutch preschool playgroups that are attended only for a few short sessions over a maximum of two years, and the early childhood education approaches that are applied, cover only part of the total childcare time. More and longer childcare sessions with early childhood education approaches may reveal stronger evidence.

On the other hand, the studied childcare quality characteristics may be too weak proxies of the actual quality of childcare that children experience. All quality characteristics studied in this dissertation are generated by averaging of quality across all children of a specific childcare group or even across all groups in one centre. The quality variation for individual children is thus not fully covered, which may introduce measurement errors. Unobserved sorting within childcare groups and centres and selective behaviour of childcare staff may increase the quality variation across children.

Also, the studied quality characteristics may not represent those factors that are actually most effective in influencing child development. The quality indicators studied in this dissertation represent aspects of childcare that are measurable by policymakers and thus can be influenced through large-scale policies. Factors that are of a 'softer' character, as for example the behaviour of childcare staff to individual children or the atmosphere in the centre, cannot be measured easily, despite the fact that they are likely to be of much importance for child development.

And investments into childcare quality may have not matured to the full degree yet within the studied period, which may be particularly the case given the recent initiatives to increase quality investments. A number of such initiatives in both the Netherlands and Denmark during the studied time periods may have caused a bigger variation in the progress of improving quality characteristics, but they may not have shown their full impact yet. For example, if the quality indicator reflects application of an early childhood education approach in a childcare centre, the quality initiative may not have been fully implemented yet. This may have been the case for the studied child-centred education programme *Speelplezier*, which was in its piloting stage during the studied period. Funds to train staff may still be too limited, institutions may need time to adapt, not all staff may have been trained in the particular method yet, or not all groups in a childcare centre may have introduced it yet.

Nap-Kolhoff et al. (2008) argue that the quality of services is still insufficiently documented to distinguish beneficial, high-quality investments from non-

beneficial, low-quality investments. Durlak (2010) points out similar problems in the international evaluation literature on early childhood education investments, stating that the implementation of a programme can never be completely measured and that it is up to the researcher to make a meaningful choice of measurement.

Initiatives to improve the quality of childcare investments are very costly. Small-scale, intensive childcare investments have delivered the expected positive results, and examples of such investments provided evidence that is used in favour of larger scale investments (see e.g. Heckman & Masterov, 2007). However, reaching the same high quality at a universal scale is much more costly, demanding that countries devote much larger shares of their gross domestic product to the early childhood period. And bigger investments in selected quality aspects may not be sufficient to generate the expected gains, if complementary investments in the general well-being of the child, such as investments to improve the family and home environment, including parental leave and family health care, remain insufficient. Also, further research may be needed to identify whether early development gains fade because of insufficient follow-up investments, in particular for disadvantaged children.

Finally, the lack of results may be due to the empirical strategy chosen. If estimation models assume linear relationships between explanatory variables and child outcomes while there are actually non-linear correlations, effects may be averaged out. An investigation into this issue in chapter II, using quantile regression, indicates that there may be indeed a non-linear relationship between childcare spells and child outcomes. Similarly, the estimates may be very sensitive to outliers, and effects may vary across the distribution. Quantile estimates in all three studies have indicated that estimations at the median may hide the impact of childcare quality to some degree. Studies of effects across subgroups, e.g. by applying interaction terms with the studied quality indicators or by comparing estimates of different sub-samples, indicate that there is some heterogeneity of effects. However, the studied subgroups may be too broad or consist of numbers too small to come to significant estimates.

Also, there may be selection mechanisms that work in the opposite direction, e.g. if better children sort into better quality centres, effects may to some degree remain spurious. To assess potential biases from unobserved heterogeneity, I have applied a number of alternative identification strategies to see whether effect estimates are consistent. As ordinary least square estimates, those alternative strategies may have their own weaknesses. For example, instrumental variables that are applied to introduce an element of randomness on the treatment allocation may predict the treatment only at a certain range - instrumental variable estimates reflecting only local average treatment effects.

The optimal research design would be provided by a controlled social experiment that assures a (quasi)random allocation of different qualities, perhaps at the municipal (or county) level – high enough to reduce the risk that parents vote with their feet by putting their to another childcare arrangement. This could be reached, by allocating subsidies for different quality aspects across different municipalities. The allocation of different qualities would not pose ethical questions as long as there is contradicting evidence about the impact of different quality characteristics on child development. Countries like Denmark, where extensive administrative data is available, may provide sufficient data to evaluate such experiments on the large-scale - if the allocation of subsidies would be administered, too. However, problematic would be to assure a random allocation of subsidies for different quality aspects without any political bias in the allocation process. And, to assure that municipalities keep participating in the experiment, a clear strategy would be needed on what to do with municipalities that received subsidies for less relevant quality aspects and how successful quality aspects could be realised across all municipalities.

Finally, the scarce evidence might imply that there are no significant effects of childcare quality aspects on child development whatsoever, implying that the returns to high-quality early childhood investments as compared to high-quality investments at later ages may be overstated (Heckman et al., 2002).

3. Policy implications

Policymakers need objectively measurable and comparable quality indicators to inform their decisions about investing in early childhood care and education. The growing public investments in childcare provisions are increasingly a subject of debate on whether public resources are being spent effectively. The three studies provide policymakers with new evidence on aspects of childcare quality that they can directly influence through the public resources that they allocate to the improvement of childcare quality, e.g. via subsidies to train childcare staff in specific early childhood education approaches, as well as the quality standards and conditions that they link to the allocation of such resources, such as the number of staff per child.

The implications for the policymaker are less straightforward. There is large-scale theoretical support for investments into better childcare quality, and such investments are often seen as essential to justify the extension of childcare provisions to allow maternal employment. However, as indicated in the previous section, the weak evidence on the expected positive returns from investing in better quality of early childhood care and education does not necessarily imply that quality does really matter. The shortage of evidence puts a large question

mark on whether policymakers can reach the desired social returns and reduce inequality of chances in life by extending public investments in childcare.

Like its Scandinavian neighbours, Denmark is a leading example of a country investing substantial public resources into high-quality, near universal childcare provisions. Other countries use Denmark as an example for positive outcomes of such investments, namely that the inequality in child outcomes can be reduced. However, a more comprehensive look reveals that those investments are complemented by a whole range of other public investments to reduce inequality. This dissertation cautions policymakers of other countries about the importance of taking such complementary policies and their effects on child outcomes into account. Looking at childcare quality aspects might be a too narrow a focus when searching for the most effective policies.

Nonetheless, the evidence provided in this dissertation does not provide any ground for those who warn the public against extending childcare investments. None of the three studies provides any reason to believe that, for example, the extension of the Dutch childcare provisions would eventually lead to harmful effects for children. The evidence for Denmark indicates that even a more full-time system may provide positive rather than negative results.

Despite the lack of evidence on quality aspects, policymakers are likely to face growing discussions of childcare policies. Many countries are working to meet the demand for childcare places. However, along with the introduction of guaranteed places or lowering of compulsory schooling ages there are new questions arising about the best way to invest public resources. The discussions in Denmark in the late 1990s point out a direction other countries may follow. Denmark's lower inequality in child outcomes and growing birth rates may indicate that its childcare and family policies may at this stage better serve the needs of children and families than other countries' such as, for example, Germany's, where there is intensive public debate about the best policies to assure sufficient childcare provisions to serve the high demand. Childcare investments may be a way to consolidate mothers' work-related needs with children's needs for additional development stimuli, in particular if they come from disadvantaged backgrounds.

Policymakers should also pay attention to the quality of parents' early childhood caregiving. Recent trends show that despite growing attendance rates for non-parental childcare, parents are tending to spend more development-oriented time with their children. Policies that can support such quality improvement in parental childcare may turn out to be more effective than quality initiatives in non-parental childcare provisions, especially when children are still spending only relatively short spells at the latter. Perhaps childcare institutions should take up the general care burden (optimally at good quality), while parents (even where both are working) start investing much more high-quality time with their children.

4. Limitations

The results of this dissertation are based on rich datasets and verified under application of valid empirical strategies. However, the nature of the data we used and the empirical strategies we applied resulted in some limitations that constrain the interpretation of those results.

First, the interpretation of results neglects possible equilibrium effects, e.g. what would happen to the quality and price of childcare provisions if all parents choose high-quality provisions? Would the funds needed to provide universal high-quality childcare be removed elsewhere, where they might have much more substantial effects on child development, perhaps creating negative effects that cancel out childcare gains? While there is a general trend to extend childcare provisions through investment of bigger shares of public and private budgets, those investments have tended to prioritize increasing the quantity rather than the quality of childcare provisions. Reallocating funds to focus on quality rather than quantity would not necessarily result in better child outcomes, as disadvantaged children might be excluded from high-quality provisions. Also, initiatives to boost the quality of childcare provisions may imply that fewer resources are available for later human resource investments. Initial gains from higher quality might then fade out, in particular if disadvantaged children experience lower quality at later education levels.

The three studies are also limited in being static rather than dynamic assessments of the impact of childcare quality. Static analysis may not be able to reveal effects of quality characteristics that have an impact in a dynamic manner, e.g. stimulating development with interactions of different development domains and promoting complementary and multiplying effects across later ages. Having more child outcomes at different ages – before and after childcare attendance – and across different development domains available will allow a deeper analysis. Richer longitudinal studies with multiple observations at different ages may allow a more dynamic assessment of the impact of childcare quality characteristics on child development. In particular, the newly generated datasets for the second and third study set the ground for future studies to look at outcomes of more child development domains. The Danish administrative records, which have been merged for the third study, are censored to left in the child registries, i.e. the exact enrolment years before 1995 are not reported. An extension to multiple cohorts will allow a stronger identification that incorporate preschool fixed effects.

The children that are covered by the *Moelejaan* dataset, which is used for the second study, are followed through their later schooling; tapping further child outcomes from school administrations has already been scheduled. First results are available for children who benefited from the quasi-experimental extension of early childhood education in South Limburg in recent years (for preliminary

evidence, see Jungbluth et al., 2011). Also, first labour market records are soon available on the studied children who attended Danish preschools in the early 1990s. Besides assessing wider child development domains at different ages, future research on the effects of childcare quality may also benefit from richer options to assess children's initial endowments before childcare attendance, i.e. the ex-ante family background, maternal health during pregnancy, early childhood health, genetic predispositions as well as early childhood development assessments, as far as they are reliable in identifying such endowments.⁶¹ Until such information is available, the investigation of sibling fixed effects may be a promising route, as long as sufficient sibling variation exists in the data; unfortunately, the studied datasets do not provide sufficient variation in this regard.

Two of the three studies face also some limitation with respect to their external validity. The first study allows a comparison between those children who are cared for at home and those who are cared for at a childcare centre. However, the other two studies only look at the group of 'insiders': those who are already attending centre-based childcare. While the share of children not attending centre-based childcare in Denmark is rather small, in the Netherlands it is still rather substantial. Quality effects may not necessarily translate to those children in the same way. As those children who are not in centre-based childcare are likely to have more disadvantaged backgrounds, we may expect that the returns from high-quality childcare experience may be higher, though. And additional research is necessary to understand the factors that make parents choose home- over centre-based childcare options.

Furthermore, the external validity is limited to the degree that childcare arrangements vary between countries with many idiosyncratic differences, e.g. with respect to the family and labour market structures, pedagogical philosophies and the childcare industry itself. Even though the internal validity of conclusions of the three studies in this dissertation shall be guaranteed by choosing adequate identification strategies, policy implications must be converted to other country settings with some care. Without a directly comparable dataset across countries such as, for example, a PISA study and strong isolation of a hypothesis that can be compared across countries, external validity remains limited. However, the questions put forward in this dissertation and conclusions drawn upon them, as well as the applied methods and used datasets, can serve as a rich input for researchers studying other country cases.

⁶¹ For an application of initial assessments for the *Moelejaan* research, see section 1 of Jungbluth et al. (2011).

And finally, the three studies are limited in the degree to which they can unveil patterns of heterogeneity of quality effects. Adding interaction terms for children of single, less educated, poor and ethnic minority parents as well as boys has revealed some evidence on such heterogeneity but produced for most of the disadvantaged sub-samples no significant differences versus the main samples. Quantile regressions have tested effects across the outcome distribution, finding some indication for stronger effects of quality on the lower end of the distribution. Overall, more research is needed into specific subgroups to reveal the full scope of heterogeneity of effects. Such research may, for example, assess the representativeness of sub-samples and then study those sub-samples separately.

5. Concluding remarks

The massive extensions of childcare provisions in recent years have raised a number of worries whether children may benefit or be harmed by being raised through part of their early childhood in childcare centres. Given the sensitivity of early childhood in particular for cognitive and language development, parents and policymakers have increasingly questioned the impact of such centre-based childcare provisions. The little data that has been collected through the early childhood and international literature has remained incomplete, so that many policy discussions remain inconclusive, e.g. on how to best allocate public resources to stimulate child development.

The long, unresolved discussion on whether child development is merely a result of nature or nurture has increased such inconclusiveness. The unsettled discussions reflect to some degree also the long debate on the true impact of school investments versus direct investments into families (see, for example, Coleman, 1966; Jencks, 1979; Mayer, 1997). Yet, the literature moves towards providing more convincing answers. Access to rich new datasets and innovative empirical methods are providing new answers to help resolve earlier debates. The studies of this dissertation contribute to this resolution process.

The hypotheses tested in this dissertation reflect the high hopes that are put in early childhood care and education: specifically, that sufficient quality may allow children to develop just as well as when they are exclusively taken care of by their parents, or even better. In particular, disadvantaged children are expected to benefit from high-quality childcare experiences. Public investments in centre-based care may thus not only permit mothers to work but also to stimulate the development of their children in line with the policy goal of promoting equal opportunity. Policies to invest in early childhood development are expected to be a sound investment; they are argued to be perfectly Paretian, maximizing efficiency and equity at the same time (in line with, for example, Cunha & Heckman, 2007). Even though the share of public resources in childcare provisions is still rather

small, policymakers hope to gain high social returns, through increasing children's skill attainment at later stages (self-productivity) and by facilitating the productivity of later skill investments (complementarity).

The new evidence added through this dissertation looks at a first glance as if all the high hopes in childcare investments have not been justified, because too many effect estimates are too small in scale or insignificant. But, as discussed before, there are a number of possible reasons why effects may not be identifiable. Nonetheless, given the evidence available thus far, for now we need to argue for investing in childcare mainly to facilitate maternal participation in the labour market. To a large degree, investments into childcare quality can be based only on theoretical or piecemeal evidence. Despite a strong and widespread belief that quality matters, we cannot yet be certain that public resources are being effectively spent.

References

- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91(434), 444-455.
- Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, Part 1, pp. 1277-1366): Elsevier.
- Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, 15(4), 69-85.
- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey, US: Princeton University Press.
- Baker, M., Gruber, J., & Milligan, K. (2008). Universal Child Care, Maternal Labor Supply, and Family Well-Being. *Journal of Political Economy*, 116(4), 709-745.
- Barnett, W. S. (1992). Benefits of Compensatory Preschool Education. *Journal of Human Resources*, 27(2), 279-312.
- Barnett, W. S. (1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. *The Future of Children*, 5(3), 25-50.
- Barnett, W. S. (2008). *Preschool education and its lasting effects: Research and policy implications*. Boulder & Tempe: Education and the Public Interest Center & Education Policy Research Unit.
- Barnett, W. S. (2011). Effectiveness of Early Educational Intervention. *Science*, 333(6045), 975-977.
- Barnett, W. S., & Lamy, C. E. (2006). Estimated Impacts of Number of Years of Preschool Attendance on Vocabulary, Literacy and Math Skills at Kindergarten Entry. *NIEER Working Paper*.
- Bauchmüller, R. (2008). Evaluating causal effects of Early Childhood Care and Education Investments: A discussion of the researcher's toolkit. *MGSOG Working Paper Series*, 2008(6).

- Bauchmüller, R., & Jungbluth, P. (2008). *Technical Report of data collection for evaluation of ECE programme 'Speelplezier' in Maastricht*. Maastricht: Maastricht University, SJOANS Research Centre.
- Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and GMM: Estimation and Testing. *The Stata Journal*, 3(1), 1–31.
- Behrman, J. R., & Birdsall, N. (1983). The Quality of Schooling: Quantity Alone is Misleading. *The American Economic Review*, 73 (5), 928-946.
- Belsky, J., & Rovine, M. J. (1988). Nonmaternal Care in the First Year of Life and the Security of Infant-Parent Attachment. *Child Development*, 59, 157-167.
- Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, 93(1-2), 219-234.
- Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, 92(5), 1416-1440.
- Bernal, R. (2008). The Effect of Maternal Employment and Child Care on Children's Cognitive Development. *International Economic Review*, 49(4), 1173-1209.
- Bernal, R., & Keane, M. P. (2006). Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers. *IPR Working Paper*, 09.
- Bernal, R., & Keane, M. P. (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*, 156(1).
- Bianchi, S. M. (2000). Maternal employment and time with children: Dramatic change or surprising continuity? *Demography*, 37(4), 401-414.
- Björklund, A., & Jäntti, M. (2008). Intergenerational income mobility and the role of family background. In W. Salverda, B. Nolan & T. Smeeding (Eds.), *Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press.
- Björklund, A., & Salvanes, K. G. (2011). Education and Family Background: Mechanisms and Policies. In E. A. Hanushek, S. J. Machin & L. Wößmann (Eds.), *Handbook of the Economics of Education* (1 ed.).

Blau, D. M. (1991). The quality of child care: An economic perspective In D. M. Blau (Ed.), *The economics of child care* (pp. 145-173). New York: Russell Sage Foundation.

Blau, D. M. (1999). The Effect of Child Care Characteristics on Child Development. *Journal of Human Resources*, 34(4), 786-822.

Blau, D. M., & Hagy, A. P. (1998). The Demand for Quality in Child Care. *Journal of Political Economy*, 106(1).

Blok, H., & Leseman, P. P. M. (1996). Effecten van voorschoolse stimuleringsprogramma's: een review van reviews (transl.: Effects of preschool stimulation programmes: a review of reviews). *Pedagogische Studiën*, 73, 184-197.

Bolt, L. v. d., & Schonewille, B. (2006). *Evaluatie KLLOS programma Maastricht: Eindverslag - Procevaluatie en materiaalanalyse*. Utrecht: Sardes.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Weel, B. t. (2008). The Economics and Psychology of Personality Traits. *Journal of Human Resources*, 43(4), 972-1059.

Bowlby, J. (1969). Attachment and loss. In *Attachment* (Vol. 1). New York: Basic Books; & Hogarth Press.

Bowles, S., & Gintis, H. (2002). The Inheritance of Inequality. *Journal of Economic Perspectives*, 16(3, Summer), 3-30.

Bowles, S., Gintis, H., & Osborne Groves, M. (2005). *Unequal chances: family background and economic success*. Princeton, N.J.: Princeton University Press.

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2011). Inequality during the Early Years: Child Outcomes and Readiness to Learn in Australia, Canada, United Kingdom, and United States. *IZA Discussion Paper*, 6120.

Braga, M., Checchi, D., & Meschi, E. (2011). Institutional Reforms and Educational Attainment in Europe: A Long Run Perspective. *IZA Discussion Paper*, 6190.

Brilli, Y., Del Boca, D., & Pronzato, C. (2011). Exploring the Impacts of Public Childcare on Mothers and Children in Italy: Does Rationing Play a Role? *IZA Discussion Paper*, 5918.

Bryson, A., Dorsett, R., & Purdon, S. (2002). *The Use Of Propensity Score Matching In The Evaluation Of Active Labour Market Policies*: Policy Studies Institute and National Centre for Social Research.

Burchinal, M. R., Vandergrift, R., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, 25, 166-176.

Caliendo, M., & Kopeinig, S. (2005). Some Practical Guidance for the Implementation of Propensity Score Matching. *IZA Discussion Paper*, 1588.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development. *Teachers College Record*, 112(3), 579-620.

Campbell, F. A., Helms, R., Sparling, J. J., & Ramey, C. T. (1998). Early Childhood Programs and Success in School: The Abecedarian Study In W. S. Barnett & S. S. Boocock (Eds.), *Early Care and Education for Children in Poverty: Promises, Programs, and Long-Term Results*. Albany, NY: State University of New York Press.

Campbell, F. A., Wasik, B. H., Pungello, E., Burchinal, M., Barbarin, O., Kainz, K., et al. (2008). Young Adult Outcomes of the Abecedarian and CARE Early Childhood Educational Interventions. *Early Childhood Research Quarterly*, 23(4), 452-466.

Card, D., & Krueger, A. B. (1996). Labor Market Effects of School Quality: Theory and Evidence. *NBER Working Paper*, 5450.

CBS Statline. (2009, May 13). Bevolking Zuid-Limburg: jongen en meisjes van leeftijd 4-6 per 1 januari 2009 (transl.: population in South-Limburg: boys and girls at the ages 4-6 by January 1, 2009). from statline.cbs.nl

CBS Statline. (2010, June 7). 6 jarigen in (speciaal) basisonderwijs in de COROP regio Zuid-Limburg, 2009/10 (transl.: 6-year old in (special) elementary education in the COROP region of South-Limubrg, 2009/10). from statline.cbs.nl

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Whitmore Schanzenbach, D., & Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR. *Quarterly Journal of Economics*, 126(4), 1593-1660.

Cleveland, G., & Krashinsky, M. (2003). *Financing ECEC services in OECD countries*. Scarborough: University of Toronto.

Clewell, B. C., & Villegas, A. M. (1998). Diversifying the Teaching Force to Improve Urban Schools: Meeting the Challenge - Introduction. *Education and Urban Society*, 31(1), 3-17.

Cochran, W. G. (1963). *Sampling techniques* (2nd ed.). New York: John Wiley and Sons.

Coleman, J. S. (1966). *Equality of Educational Opportunity (COLEMAN) Study (EEOS)*. Washington, DC: United States Department of Education.

Connolly, S., Micklewright, J., & Nickell, S. (1992). The Occupational Success of Young Men Who Left School at Sixteen. *Oxford Economic Papers*, 44(3), 460-479.

Cook, R. D. (1998). *Regression graphics: ideas for studying regressions through graphics*: John Wiley & Sons.

Copple, C., & Bredekamp, S. (1997). *Developmentally Appropriate Practice in Early Childhood Programs Serving Children from Birth through Age 8* (3rd ed.). Washington, DC: National Association for the Education of You (NAEYC).

Copple, C., & Bredekamp, S. (2009). *Developmentally Appropriate Practice in Early Childhood Programs Serving Children from Birth through Age 8* (3rd ed.). Washington, DC: National Association for the Education of Young Children (NAEYC).

Cox, H., Jungbluth, P., & Rodigas, E. (2008). *Pilot Voor- en Vroegschoolse Educatie Zuid-Limburg - "Moelejaan": een inhaalslag richting sterke startkansen*. Kerkrade: VVE-pilot Zuid Limburg "Moelejaan".

Cunha, F., & Heckman, J. J. (2007). The Technology of Skill Formation. *The American Economic Review*, 97(2), 31-47.

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4), 738-782.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 1, pp. 697-812). Amsterdam: Elsevier.

Currie, J. (1998). The Effect of Welfare on Child Outcomes: What We Know and What We Need to Know. *JCPR Working Paper*, 26.

Currie, J. (2000). Early Childhood Intervention Programs: What Do We Know? *JCPR Working Paper*, 169, 213-238.

Currie, J., & Thomas, D. (1995). Does Head Start Make a Difference? *The American Economic Review*, 85(3), 341-364.

Currie, J., & Thomas, D. (2001). Early Test Scores, School Quality and SES: Longrun Effects on Wage and Employment Outcomes. *Worker Wellbeing in a Changing Labor Market*, 20, 103-132.

Datta Gupta, N., & Simonsen, M. (2010). Non-Cognitive Child Outcomes and Universal High Quality Child Care. *Journal of Public Economics*, 94(1-2), 30-43.

Datta Gupta, N., & Simonsen, M. (2011). Where to Put the Kids? Effects of Type of Non-parental Child Care on Pre-teen Skills and Risky Behavior. *IZA Discussion Paper*, 5848.

Deding, M. C., & Filges, T. (2004). Arbejdsløs og geografisk mobil? (transl.: Unemployed and geographically mobile?). In N. H. Bjørn (Ed.), *Bolig, mobilitet og marginalisering på arbejdsmarkedet* (transl.: *Residence, Mobility and Marginalization on the Labor Market*) (pp. 86-115): The Danish National Institute of Social Research.

Dee, T. S. (2004). Teachers, Race and Student Achievement in a Randomized Experiment. *Review of Economics and Statistics*, 86(1), 195-210.

Dee, T. S. (2005a). A Teacher Like Me: Does Race, Ethnicity or Gender Matter? *The American Economic Review*, 95(2).

Dee, T. S. (2005b). Teachers and the Gender Gaps in Student Achievement. *NBER Working Paper*, 11660.

Dee, T. S. (2006). The Why Chromosome: How a teacher's gender affects boys and girls. *Education Next*, 6(4), 69-75.

Dekker, J., de Fijter, R., & Veen, A. (2000). *Keuzegids VVE. Programma's - deskundigheidsbevordering en instrumenten op het gebied van Voor- en Vroegschoolse Educatie* (transl.: *Guide to various choices - knowledge stimulation and instruments in the area of early childhood and preschool education*). Den Haag/Utrecht: Makelaar VVE.

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3), 111-134.

Dickens, W. T., Sawhill, I. V., & Tebbs, J. (2006). The Effects of Investing in Early Education on Economic Growth. *The Brookings Institution: Policy Brief*, 153.

Driessen, G. (2004). A large-scale longitudinal study of the utilization and effects of early childhood education and care in The Netherlands. *Early Child Development and Care*, 174(7), 667-689.

Driessen, G., & Doesborgh, J. (2003). *Voor- en Vroegschoolse Educatie en cognitieve en niet-cognitieve competenties van jonge kinderen*. Nijmegen: Instituut voor Toegepaste Sociale wetenschappen (ITS).

Driessen, G., Van Langen, A., & Vierke, H. (2006). *Basisonderwijs: veldwerkverslag, leerlinggegevens en oudervragenlijsten - Basisrapportage PRIMA-cohortonderzoek - Zesde meting 2004/05*. Nijmegen: ITS.

Dronkers, J. (2010). Positive but also negative effects of ethnic diversity in schools on educational performance? An empirical test using cross-national PISA data. *MPRA Paper*, 25598.

Dumas, C., & Lefranc, A. (2010). Early schooling and later outcomes : Evidence from pre-school extension in France. *Thema Working Paper*, 7.

Duncan, G. J. (2003). Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development. *Child Development*, 74(5), 1454-1475.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K. A., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428-1446.

Duncan, G. J., Magnuson, K. A., & Ludwig, J. (2004). The Endogeneity Problem in Developmental Studies. *Research in Human Development*, 1(1), 59-80.

Durlak, J. A. (2010). The importance of doing well in whatever you do: A commentary on the special section, "Implementation research in early childhood education". *Early Childhood Research Quarterly*, 25(3), 348-357.

Dustmann, C., & Preston, I. (2001). Attitudes to ethnic minorities, ethnic context and location decisions. *Economic Journal*, 111(470), 353-373.

Dustmann, C., Puhani, P. A., & Schönberg, U. (2012). The Long-term Effects of School Quality on Labor Market Outcomes and Educational Attainment. *CReAM Discussion Paper*, 8.

EC COM. (2008). *An updated strategic framework for EU cooperation in education and training* (No. 865). Brussels.

Ermisch, J. F. (1989). Purchased Child Care, Optimal Family Size and Mother's Employment: Theory and Econometric Analysis. *Journal of Population Economics*, 2(2), 79-102.

Esping-Andersen, G. (2004). Unequal opportunities and the mechanisms of social inheritance. In M. Corak (Ed.), *Generational Income Mobility in North America and Europe*: Cambridge University Press.

Esping-Andersen, G. (2005). Children in the welfare state. A social investment approach. *DemoSoc working paper*, 10, 1-33.

Esping-Andersen, G. (2006). Families, Government and the Distribution of Skills: Universitat Pompeu Fabra.

Esping-Andersen, G. (2008). Childhood Investments and Skill Formation. *International Tax and Public Finance*, 15(1), 19-44.

Esping-Andersen, G., Garfinkel, I., Han, W.-J., Magnuson, K., Wagner, S., & Waldfogel, J. (2011). Child care and school performance in Denmark and the United States. *Children and Youth Services Review*, 34(3), 576-589.

Eurydice. (2007). *Organisatie van het onderwijssysteem in Nederland (transl.: Organisation of the education system in the Netherlands)* 2006/07. Brussels: EC, DG-EAC.

Felfe, C., & Lalive, R. (2010). How Does Early Childcare affect Child Development? Learning from the Children of German Unification. University of Lausanne.

Fitzpatrick, M. D. (2008). Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement. *B.E. Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy*, 8(1).

Fletcher, J. M. (2012). The Effects of Personality Traits on Adult Labor Market Outcomes: Evidence from Siblings. *IZA Discussion Paper*, 6391.

Frank Porter Graham - Child Development Institute. (2005). Environment Rating Scales. Retrieved 9 October, 2009, from <http://www.fpg.unc.edu/>

Frede, E. C. (1998). Preschool program quality in programs for children in poverty. In W. S. Barnett & S. S. Boocock (Eds.), *Early Care and Education for Children in Poverty: Promises, Programs, and Long-Term Results* (pp. 77-98). Albany, NY: State University of New York Press.

Fredriksson, P., & Öckert, B. (2008). Resources and Student Achievement - Evidence from a Swedish Policy Reform. *The Scandinavian Journal of Economics*, 110(2), 277-296.

- Frost, C., & Thompson, S. G. (2000). Correcting for regression dilution bias: comparison of methods for a single predictor variable. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 163(2), 173-189.
- Garces, E., Thomas, D., & Currie, J. (2002). Longer-Term Effects of Head Start. *The American Economic Review*, 92(4), 999-1012.
- Gathmann, C., & Sass, B. (2012). Taxing Childcare: Effects on Family Labor Supply and Children IZA Discussion Paper, 6440.
- Gibbons, S., Machin, S. J., & Silva, O. (2006). Choice, Competition and Pupil Achievement. IZA Discussion Paper, 2214.
- Glavind, N. (2007). Guaranteed Daycare Places (GAPS), DNK 1994-2006, unpublished database. Værløse: Bureau2000.
- Gorey, K. M. (2001). Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. *School Psychology Quarterly*, 16(1), 9-30.
- Gormley, W. T., & Gayer, T. (2005). Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program. *Journal of Human Resources*, 40(3), 533-558.
- Gørtz, M., & Andersson, E. (2010). Child-to-Teacher Ratio in Daycare and Teacher Sickness Absenteeism. *AKF Working Paper*.
- Graue, E., Clements, M. A., Reynolds, A. J., & Niles, M. D. (2004). More than teacher directed or child initiated: Preschool curriculum type, parent involvement, and children's outcomes in the child-parent centers. *Education Policy Analysis Archives*, 12(72).
- Gregg, P., Washbrook, E., & ALSPAC Study team. (2003). The Effects of Early Maternal Employment on Child Development in the UK. *CMPO Working Paper Series*, 3(70).
- Gregg, P., Washbrook, E., Propper, C., & Burgess, S. (2005). The Effects of a Mother's Return to Work Decision on Child Development in the UK. *Economic Journal*, 115(501), F48-F80.
- Grilli, L., & Rampichini, C. (2011). *Propensity scores for the estimation of average treatment effects in observational studies*. Paper presented at the Training Sessions on Causal Inference, Bristol, June 28-29, 2011.

- Hanushek, E. A. (2010). The economic value of higher teacher quality. *NBER Working Paper*, 16606.
- Harrod, L. A., & Lesser, V. (2006). *The Use of Propensity Scores to Adjust for Nonignorable Nonresponse Bias*. Paper presented at the ASA Section on Survey Research Methods.
- Harvey Arnold, D., McWilliams, L., & Harvey Arnold, E. (1998). Teacher Discipline and Child Misbehavior in Day Care: Untangling Causality With Correlational Data. *Developmental Psychology*, 34(2), 276-287.
- Haskins, R. (1985). Public School Aggression among Children with Varying Day-Care Experience. *Child Development*, 56(3), 689-703.
- Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251-1271.
- Haveman, R., & Wolfe, B. (1995). The Determinants of Children's Attainments: A Review of Methods and Findings. *Journal of Economic Literature*, 33(4), 1829-1878.
- Havnes, T., & Mogstad, M. (2010). Is Universal Child Care Leveling the Playing Field? Evidence from Non-Linear Difference-in-Differences. *IZA Discussion Paper*, 4978.
- Havnes, T., & Mogstad, M. (2011). No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes. *American Economic Journal: Economic Policy*, 3(2).
- Heckman, J. J. (1976). The common structure of statistical models for truncation, sample selection and limited dependent variables and a simple estimator for such models. In *Annals of Economic and Social Measurement* (Vol. 5, pp. 475-492). New York.
- Heckman, J. J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica*, 46(4), 931-959.
- Heckman, J. J. (2008). Schools, Skills, and Synapses. *Economic Inquiry*, 46(3), 289-324.
- Heckman, J. J. (2011). Integrating Personality Psychology into Economics. *IZA Discussion Paper*, 5950.

Heckman, J. J., Ichimura, H., Smith, J., & Todd, P. E. (1998). Characterizing Selection Bias Using Experimental Data. *NBER Working Paper*, 6699.

Heckman, J. J., Ichimura, H., & Todd, J. J. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies*, 64, 605-654.

Heckman, J. J., Krueger, A. B., & Friedman, B. M. (2002). *Inequality in America: what role for human capital policies?* Cambridge, MA: MIT Press.

Heckman, J. J., Lalonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, Part 1, 1999, pp. 1865-2097): Elsevier.

Heckman, J. J., Layne-Farrar, A., & Todd, P. E. (1995). Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship. *NBER Working Paper*, 5274.

Heckman, J. J., & Macurdy, T. (1986). Labor econometrics. In J. J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics* (Vol. 3, pp. 1917-1977): Elsevier.

Heckman, J. J., & Masterov, D. V. (2007). The Productivity Argument for Investing in Young Children. *NBER Working Paper*, 13016.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yaditz, A. (2010). The Rate of Return to the High/Scope Perry Preschool Program. *Journal of Public Economics*, 94(1), 114-128.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3), 411-482.

Hohmann, M., Banet, B., & Weikart, D. P. (1979). *Young children in action*. Ypsilanti, Michigan: High/Scope Press.

Huang, F. L., Invernizzi, M. A., & Drake, E. A. (2012). The differential effects of preschool: Evidence from Virginia. *Early Childhood Research Quarterly*, 33-45.

Huerta, M. d. C., Adema, W., Baxter, J., Corak, M., Deding, M. C., Gray, M., et al. (2011). Early Maternal Employment and Child Development in Five OECD Countries. *OECD Social, Employment and Migration Working Papers*, 118.

Huffman, L. R., & Speer, P. W. (2000). Academic performance among at-risk children: The role of developmentally appropriate practices. *Early Childhood Research Quarterly*, 15(2), 167-184

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), 467-475.

Imbens, G. W., & Rubin, D. B. (2008). Causal Inference in Statistics, and in the Social and Biomedical Sciences. Unpublished [forthcoming Cambridge University Press: New York].

Imbens, G. W., & Wooldridge, J. M. (2007). Lecture 5: Instrumental Variables with Treatment Effect Heterogeneity: Local Average Treatment Effects. *Summer Institute: What's New in Econometrics* Retrieved December 5, 2011

Inspectie van het Onderwijs. (2006). *Toetsingskader voor de pilot 'Toezicht VVE in de G4'*. Utrecht: Inspectie van het Onderwijs.

Jencks, C. (1979). *Who gets ahead? The determinants of economic success in America*. New York: Basic Books.

Jensen, B. (2009). *Effekter af indsatser for socialt udsatte børn i daginstitutioner - HPA-Projektet (transl.: Effects of interventions for socially vulnerable children in day care centers - HPA-project)* (1 ed.). Copenhagen: DPU Press.

Jepma, I., Van der Vegt, A. L., & Kooiman, P. (2007). *Landelijke Monitor Voor- en Vroegsschoolse Educatie 2007*: Sardes.

Jungbluth, P. (2009). *"Moelejaan": VVE-pilot Zuid Limburg De regiobrede 'nulmeting' als start van de evaluatie*. Maastricht: Maastricht University, SJOANS Research Centre.

Jungbluth, P., Nap-Kolhoff, E., & Rodigas, E. (2011). *Peuters en kleuters in het Zuid-Limburgse onderwijs - De (voorlopige) evaluatie van de Zuid-Limburgse VVE-pilot "Moelejaan" 2009-2010*. Maastricht: Maastricht University, KAANS Research Centre.

Jungbluth, P., Rodigas, E., & Bauchmüller, R. (2009). *"Moelejaan" Baseline Inventory of Early Childhood and Elementary Education in South Limburg (data collection)*: Maastricht University, SJOANS Research Centre.

Karoly, L. A., Kilburn, M. R., & Cannon, J. S. (2005). *Early Childhood Interventions: Proven Results, Future Promise*. Santa Monica, CA: Rand.

Keane, M. P., & Wolpin, K. I. (1997). The Career Decisions of Young Men. *Journal of Political Economy*, 105(3), 473-522.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, Neurobiological and Behavioral Perspectives on Building America's Future Workforce. *Proceedings of the National Academy of Sciences*, 103(27), 10155-10162.

Koenker, R., & Basset, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33-50.

Kramarz, F., Machin, S. J., & Ouazad, A. (2009). What Makes a Test Score? The Respective Contributions of Pupils, Schools, and Peers in Achievement in English Primary Education. *LSE CEE DP*, 102.

Landvoigt, T., Mühler, G., & Pfeiffer, F. (2007). Duration and Intensity of Kindergarten Attendance and Secondary School Track Choice. *ZEW Discussion Paper*, 7(51).

Lasser, J., & Fite, K. (2011). Universal Preschool's Promise: Success in Early Childhood and Beyond. *Early Childhood Education Journal*, 39, 169-173.

Leak, J., Duncan, G. J., Li, W., Magnuson, K., Schindler, H., & Yoshikawa, H. (2010). *Is Timing Everything? How Early Childhood Education Program Impacts Vary by Starting Age, Program Duration and Time Since the End of the Program*. Paper presented at the Association for Policy Analysis and Management meetings, November 4-6, 2010.

Leamer, E. (1983). Let's take the con out of econometrics. *The American Economic Review*, 73(1), 31-43.

Lee, V. E., & Loeb, S. (1995). Where Do Head Start Attendees End Up? One Reason Why Preschool Effects Fade Out. *Educational Evaluation and Policy Analysis*, 17(1), 62-82.

Lee, W.-S. (2006). Propensity Score Matching and Variations on the Balancing Test. Unpublished paper. Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Leseman, P. P. M. (2002). *Early Childhood Education and Care for Children from Low-income or minority backgrounds*. Paris: OECD.

Leseman, P. P. M. (2009). *Tackling Social and Cultural Inequalities through Early Childhood Education and Care in Europe*. Brussels: EURYDICE Education, Audiovisual and Culture Executive Agency.

Leseman, P. P. M., Otter, M., Blok, H., & Deckers, P. (1998). Effecten van voor- en vroegschoolse educatieve centrumprogramma's. Een meta-analyse van studies gepubliceerd tussen 1985 en 1996. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 14(3), 134-154.

Leuven, E., Lindahl, M., Oosterbeek, H., & Webbink, D. (2010). Expanding schooling opportunities for 4-year-olds. *Economics of Education Review*, 29(3), 319-328.

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Litjens, I. (2011). The effects of the 2000 Dutch ECE policy - A study using propensity score modelling in estimating average treatment effects on disadvantaged children's learning outcomes. Unpublished manuscript. Graduate School of Governance, Maastricht University.

Little, R. J. A. (1986). Survey Nonresponse Adjustments for Estimates of Means. *International Statistical Review*, 54(2), 139-157.

Loeb, S., Bridges, M., Fuller, B., Rumberger, R., & Bassok, D. (2005). How Much is Too Much? The Influence of Preschool Centers on Children's Social and Cognitive Development. *NBER Working Paper*, 11812.

Ludwig, J., & Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *Quarterly Journal of Economics*, 122(1), 159-208.

Magnuson, K. A., Lahaie, C., & Waldfogel, J. (2006). Preschool and School Readiness of Children of Immigrants. *Social Science Quarterly*, 87(5).

Magnuson, K. A., Ruhm, C. J., & Waldfogel, J. (2007). Does Prekindergarten Improve School Preparation and Performance? *Economics of Education Review*, 26(1), 33-51.

Magnuson, K. A., & Waldfogel, J. (2005). Child Care, Early Education, and Racial/Ethnic Test Score Gaps at the Beginning of School. *The Future of Children*, 15(1), 169-196.

Manski, C. F. (1993). Identification Problems in the Social Sciences. *Sociological Methodology*, 23, 1-56.

- Marcon, R. A. (1992). Differential effects of three preschool models on inner-city 4-year-olds. *Early Childhood Research Quarterly*, 7(4), 517-530.
- Mariën, S. (2007, December 11). Aantal Nederlanders in basisscholen stijgt niet langer (transl.: Number of Dutch in elementary schools does not grow anymore). *Het Belang Van Limburg*,
- Marshall, N. L. (2004). The Quality of Early Child Care and Children's Development. *Current Directions in Psychological Science*, 13(4), 165-168.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., et al. (2008). Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills. *Child Development*, 79(3), 732-749.
- Mayer, S. E. (1997). *What money can't buy: family income and children's life chances*. Cambridge, MA: Harvard University Press.
- Melhuish, E. C. (2011). Preschool Matters. *Science*, 333(6040), 299-300.
- Middleton, P. (2009). *Key Figures 2004-2008 - Education, Culture and Science*. The Hague: Dutch Ministry of Education, Culture and Science (OCW).
- Montie, J. E., Xiang, Z., & Schweinhart, L. J. (2006). Preschool experience in 10 countries: cognitive and language performance at age 7. *Early Childhood Research Quarterly*, 21(313-331).
- Nap-Kolhoff, E., & Schilt-Mol, T. v. (2008, 6-7 March). *Programmes for early childhood education in pre-school playgroups in the Netherlands: what works in practice?* Paper presented at the Second International Conference on Early Childhood Education, Arnhem.
- Nap-Kolhoff, E., Schilt-Mol, T. v., Simons, M., Sontag, L., Steensel, R. v., & Vallen, T. (2008). *VVE onder de loep - Een studie naar de uitvoering en effectiviteit van voor- en vroegschoolse educatie*. Tilburg: IVA beleidsonderzoek en advies & Babylon - Centrum voor Studies van de Multiculturele Samenleving.
- Nap-Kolhoff, E., Van Schilt-Mol, T., & Vallen, T. (2008). *VVE-catalogus. 'Handreiking voor een verantwoorde keuze en uitvoering van programma's voor voor- en vroegschoolse educatie'*. Amsterdam: Aksant.

Nores, M., & Barnett, W. S. (2009). Benefits of early childhood interventions across the world: (Under) Investing in the very young. *Economics of Education Review*, 1045.

Nores, M., Belfield, C. R., Barnett, W. S., & Schweinhart, L. J. (2005). Updating the Economic Impacts of the High/Scope Perry Preschool Program. *Educational Evaluation and Policy Analysis*, 27(3), 245-261.

OCW. (2000). Regeling Voor- en Vroegschoolse Educatie (VVE) (transl.: Regulation for Early childhood and Preschool Education). *Uitleg*, 24(12), 11-18.

OCW. (2001). *Besluit landelijk beleidskader gemeentelijk onderwijsachterstandenbeleid* (transl.: *Decision on national legislation for municipal policies to address inequality in education*) 2002-2006 - *Staatsblad* 445. 's-Gravenhage: Sdu Uitgevers.

OCW. (2006). *Besluit vaststelling doelstelling en bekostiging onderwijsachterstandenbeleid* (transl.: *Decisionmaking on targeting and financing of municipal policies to address inequality in education*) 2006-2010 - *Staatsblad* 451 's-Gravenhage: Sdu Uitgevers.

OCW. (2008). *Beleidsnota: Harmonisatie en uitbreiding VVE aanbod* (transl.: *Policy Note: Harmonisation and extension of Early Childhood and Preschool Education provisions*). The Hague.

OECD. (2000). *Background Report: Early Childhood Education and Care Policy in Denmark*. Copenhagen: The Ministry of Social Affairs in consultation with the Ministry of Education.

OECD. (2006). *Starting strong II: early childhood education and care*. Paris: OECD.

OECD. (2007). *Babies and Bosses - Reconciling Work and Family Life: A Synthesis of Findings for OECD Countries* (Complete ed. Vol. 14).

OECD. (2011). *Labour Force Statistics 1989 - 2009: 2010 Edition (E-book)*. Paris: OECD.

OECD Review Team. (1999). *Country Note: Early Childhood Education and Care Policy in the Netherlands*. Paris: OECD.

Parente, M. D. C., & Santos Silva, J. M. C. (2012). A cautionary note on tests of overidentifying restrictions. *Economics Letters*, 115, 314-317.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., et al. (2001). The Relation of Preschool Child-Care Quality to

Children's Cognitive and Social Developmental Trajectories through Second Grade. *Child Development*, 72(5), 1534-1553.

Pennings, T. (2009). Databank effectieve jeugdinterventies: beschrijving 'Speelplezier'. Retrieved 13.07.2009, from <http://www.nji.nl/jeugdinterventies>

Pfau-Effinger, B. (1999). Change of Family Policies in the Socio-Cultural Context of European Societies. In A. Leira (Ed.), *Family Policies: Practices, Policies and Values* (Vol. 18, pp. 135-159). Stamford: JAI Press.

Plug, E., & Vijverberg, W. (2001). Schooling, Family Background, and Adoption: Does Family Income Matter? *IZA Discussion Paper*, 246.

Potter, F., Grau, E., Williams, S., Diaz-Tena, N., & Lepidus Carlson, B. (2006). *An Application of Propensity Modeling: Comparing Unweighted and Weighted Logistic Regression Models for Nonresponse Adjustments*. Paper presented at the ASA Section on Survey Research Methods: Joint Statistical Meeting 2006.

Rässler, S., & Riphahn, R. (2006). Survey item nonresponse and its treatment. *AStA Advances in Statistical Analysis*, Springer, 90(1), 217-232.

Reynolds, A. J. (1993). Effects of a Preschool Plus Follow-on Intervention Program for Children at Risk. *Developmental Psychology*, 30, 787-804.

Reynolds, A. J. (2000). *Success in early intervention: The Chicago Child-Parent Centers*. Lincoln, NE: University of Nebraska Press.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2), 417-458.

Rodigas, E. (2009). *Verslag Dataverzameling Nulmeting "Moelejaan" 2008-2009 (tot versie 3b)*. Sittard-Geleen: O2-onderwijsadvies.

Rolnick, A. J., & Grunewald, R. (2003). Early Childhood Development: Economic Development with a High Public Return. *The Region*, 17(4, Supplement), 6-12.

Ronfeldt, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). How Teacher Turnover Harms Student Achievement. *NBER Working Paper*, 17176.

Rosenbaum, P., & Rubin, D. B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, 79, 516-524.

Ruhm, C. J. (2005). Maternal Employment and Adolescent Development. *Labour Economics*, 15(5), 958-983.

Ruhm, C. J., & Waldfogel, J. (2011). Long-Term Effects of Early Childhood Care and Education. *IZA Discussion Paper*, 6149.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, 26, 393-415.

Schütz, G., Ursprung, H. W., & Wößmann, L. (2008). Education Policy and Equality of Opportunity. *Kyklos International Review of Social Sciences*, 61(2), 279-308.

Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant Benefits: The High/Scope Perry Preschool Study through Age 27 (Monograph 10)*. Ypsilanti, MI: High/Scope Educational Research Foundation.

Schweinhart, L. J., & Weikart, D. P. (1981). Effects of the Perry Preschool Program on Youths Through Age 15. *Journal of Early Intervention*, 4, 29-39.

Schweinhart, L. J., Weikart, D. P., & Larner, M. B. (1986). Consequences of three preschool curriculum models through age 15. *Early Childhood Research Quarterly*, 1(1), 15-45.

Simonsen, M. (2005). Availability and price of high quality child care and female employment. *University of Aarhus Economics working paper*(8).

Smith, J., & Todd, P. E. (2005). Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators? *Journal of Econometrics*, 5(1-2), 305-353.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3), 557-586.

Statistics Denmark. (1998). *Statistisk årbog (transl.: Statistical Yearbook)*. Copenhagen.

Statistics Denmark. (2006a). Børn fordelt på søskende de bor sammen med - antal og kombination efter område, alder, familietype, antal søskende og kombination af søskende (transl.: Children by region, age, family type, number of siblings and combination of siblings). (Publication no. BRN05). Retrieved 27 August 2010, from Statbank: <http://www.statbank.dk>

Statistics Denmark. (2006b). Valg til kommunale råd efter område, parti og stemmer/kandidater/køn (transl.: Elections to municipality councils by region,

party and votes/candidates/sex). (Publication no. VALGK3X). Retrieved 27 August 2010, from Statbank: <http://www.statbank.dk>

Stipek, D. J., Feiler, R., Byler, P., Ryan, R., Milburn, S., & Salmon, J. M. (1998). Good Beginnings: What Difference Does the Program Make in Preparing Young Children for School? *Journal of Applied Developmental Psychology*, 19(1), 41-66.

Stipek, D. J., Feiler, R., Daniels, D., & Milburn, S. (1995). Effects of different instructional approaches on young children's achievement and motivation. *Child Development*, 66(1), 209-223.

Tobin, J. J., Wu, D. Y. H., & Davidson, D. H. (1991). *Preschool in Three Cultures: Japan, China and the United States*. New Haven, CT: Yale University Press.

Todd, P. E. (2008). Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated. In T. P. Schultz & J. A. Strauss (Eds.), *Handbook of Development Economics* (1 ed., Vol. 4): Elsevier B.V.

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function Cognitive Achievement. *Economic Journal*, 113(F3-F33).

UNICEF. (2000). A League Table of Child Poverty In Rich Countries. *Innocenti Report Cards*, 1.

UNICEF. (2005). Child poverty in rich countries 2005. *Innocenti Report Cards*, 6.

UNICEF. (2008). *The child care transition - A league table of early childhood education and care in economically advanced countries (Report Card 8)*. Florence: Innocenti Research Centre.

Van der Vegt, A. L., Studulski, F., & Klopogge, J. (2007). *Voorschoolse voorzieningen en onderwijskansen - review van de onderzoeksliteratuur (transl.: Preschool arrangements and opportunities in education - review of the research literature)*. Utrecht: Sardes.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., & NICHD. (2010). Do Effects of Early Child Care Extend to Age 15 Years? Results From the NICHD Study of Early Child Care and Youth Development. *Child Development*, 81(3), 737-756.

Veen, A., Roeleveld, J., & Leseman, P. P. M. (2000). *Evaluatie van Kaleidoscoop en Piramide: Eindrapportage* (No. 576). Amsterdam: SCO-Kohnstamm Instituut van de Faculteit der Maatschappij- en Gedragswetenschappen, Universiteit van Amsterdam.

- Verbeek, M. (2000). *A guide to modern econometrics*: John Wiley & Sons.
- Villegas, A. M., & Jordan Irvine, J. (2010). Diversifying the Teaching Force: An Examination of Major Arguments. *The Urban Review*, 42(3), 175-192.
- Waldfoegel, J. (2004). Child care, women's employment, and child outcomes. *Journal of Population Economics*, 15(3), 527-548.
- Waldfoegel, J., Han, W.-J., & Brooks-Gunn, J. (2002). The Effects of Early Maternal Employment on Child Cognitive Development. *Demography*, 39(2), 369-392.
- Wooldridge, J. M. (1995). Score diagnostics for linear models estimated by two stage least squares. In G. S. Madala, P. C. B. Phillips & T. N. Srinivasan (Eds.), *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C.R. Rao* (pp. 66-87). Oxford: Blackwell.
- Wößmann, L. (2004). How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States. *IZA Discussion Paper*, 1284.
- Wouterse-Schmitz, M. (2006). *Speelplezier - moduleklapper voor leidsters en leerkrachten*. Heerlen: Consent.
- Wun, L.-M., Ezzati-Rice, T. M., Baskin, R., Greenblatt, J., Zodet, M., Potter, F., et al. (2004). Using Propensity Scores to Adjust Weights to Compensate for Dwelling Unit Level Nonresponse in the Medical Expenditure Panel Survey. *MEPS Working Paper* (4004).
- Yansaneh, I. S. (2003, December 3-5). *Construction and use of sample weights*. Paper presented at the Expert Group Meeting to Review the Draft Handbook on Designing of Household Sample Surveys.

Executive Summary

Investing in Early Childhood Care and Education: The Impact of Quality on Inequality

This dissertation aims at informing policymakers about ways to invest in early childhood care and education. It adds important evidence on a number of quality characteristics that policymakers may directly influence to stimulate child development. Historically public investments in early childhood were primarily aiming at enabling maternal employment. However, policymakers have increasingly acknowledged them as instruments to stimulate social mobility of children who are at risk of falling behind at school, for example, because of growing up in poor households, with single or low-educated parents, belonging to an ethnic minority, or being a boy. The importance of mobilizing the full potential of children has been underlined by the fact that, despite economic progress, equality of opportunities has not been reached in most countries, whereas demographic trends threaten to raise the burden young generations have to carry in future. At the same time, early childhood research has supported the rationale to invest public resources at early ages to facilitate positive development dynamics at later ages, and thus to reap high returns for the society as a whole.

Governments in many countries are rapidly extending public investments in early childhood care and education; some countries already reach near-universal high-quality provisions. Important quality aspects that policymakers can influence are in particular linked with the childcare type, the educational approach applied, and the availability of adequate staffs. However, the relevance of such aspects for child development and in compensating undesired heterogeneity in child outcomes, has received less attention so far. This dissertation provides new evidence on those quality aspects from empirical analyses of three unique datasets, two of which have been made available for research for the first time. Two settings are being investigated – the Netherlands, a country where formal childcare provisions has expanded rapidly in recent years, and Denmark, a country with a long history of universal high-quality childcare.

The first study (chapter II) addresses effects of centre- versus home-based childcare; using rich survey data of a large-scale cohort study covering children growing up in the Netherlands at the beginning of the new millennium. The Netherlands has made substantial investments in the last two decades to make the market of centre-based provisions more professional and far-reaching, and to improve children's school readiness. The study analyses the impact of experiencing centre- rather than home-based childcare on a range of child outcomes at the beginning of elementary schooling. The study accounts for

possible non-linearity in the correlation between the childcare experience and child outcomes. Results of ordinary least square estimates and sensitivity analyses, i.e. instrumental variable and structural equation modelling approaches, do not support the significant short-term effects of centre-based childcare stated in the literature.

The second study (chapter III) investigates whether children do better in school if they attend a preschool using a child-centred rather than a teacher-based early education approach, or none at all. Early childhood education is expected to help in particular disadvantaged children. However, there is disagreement on whether education at that age should be more teacher- and curriculum-based, or rather centred on the development of an individual child. The study look in particular at child-centred education programmes (predominantly 'Speelplezier') applied in the South of Limburg, a Dutch province. The study provides the first analysis of a rich dataset collected in 2008/09, which covers recently enrolled children at elementary schools in that region. Different estimation techniques show that child-centred early childhood education has not yet led to the anticipated significant effects on child outcomes at early elementary schooling.

The third study (chapter IV) examines the long-run benefits from universal high-quality preschooling, in particular the role of quality aspects of the preschool staff in children's performance at the end of primary school (age 16). A unique match of Danish administrative registers permits to link data on children to their attended preschool, and to generate five main indicators describing the staff quality of those preschools (staff per child; turnover of staff; and share of professional, male and ethnic-minority staff). Controlling for child background factors, the study investigates effects of those indicators on long-term outcomes, and heterogeneity of such effects. The different analyses show modest long-run effects of better preschool staff quality on children's linguistic development. Four out of the five quality indicators are associated with significant improvements in children's test results in Danish – albeit the relationship being numerically modest. Boys benefit more from better preschool quality than girls, and ethnic minority children benefit from higher staff stability. A sensitivity analysis, an instrumental variable approach, addresses possible biases that may result from unobserved variables in the ordinary least square estimates; however, it does not produce more consistent estimates.

The results of those studies indicate that better quality may well lead to positive outcomes. However, the new evidence is not enough to justify the high hopes put in childcare investments, because too many effect estimates are too small in scale or insignificant, and the evidence on the social mobilization potential of such investments does not match the high expectations. A number of possible reasons and critical limitations to the analyses are put forward in discussion of the

evidence. As a result of what we know so far, investments in early childhood care and education have to rely primarily on the motivation to facilitate maternal participation in the labour market. Investments into specific quality aspects can be based only on theoretical or piecemeal evidence. Despite a strong and widespread belief that quality matters, we cannot yet be certain that public resources are being effectively spent.

Nederlandse Samenvatting

Investeren in Voorschoolse Opvang en Onderwijs: Het Effect van Kwaliteit op Ongelijkheid

Doel van dit proefschrift is beleidsmakers te informeren over hoe zij in voorschoolse opvang en onderwijs kunnen investeren. Het levert belangrijk bewijs over een aantal kwaliteitskenmerken die beleidsmakers rechtstreeks kunnen beïnvloeden om de ontwikkeling van kinderen te stimuleren. In het verleden waren investeringen van de overheid in de vroege kinderjaren vooral gericht op het mogelijk maken van arbeidsmarktparticipatie door moeders. Tegenwoordig erkennen beleidsmakers dit soort investeringen steeds meer als instrumenten om de sociale mobiliteit te verbeteren van kinderen die risico lopen op een onderwijsachterstand. Denk hierbij aan jongens, kinderen die opgroeien in arme huishoudens, kinderen uit eenoudergezinnen, kinderen met laagopgeleide ouders, en kinderen die behoren tot een etnische minderheid. Het belang van het mobiliseren van het volledige potentieel van kinderen wordt onderstreept door het feit dat, ondanks economische vooruitgang, in de meeste landen nog steeds geen sprake is van gelijke kansen. Bovendien dreigen demografische ontwikkelingen de toekomstige lasten voor jonge generaties te verzwaren. Tegelijkertijd ondersteunt onderzoek naar de vroege kinderjaren de redenering dat publieke middelen die al op jonge leeftijd worden geïnvesteerd een positieve ontwikkelingsdynamiek in latere jaren mogelijk maken. De samenleving kan dus een hoog rendement plukken van deze investeringen.

In veel landen zijn overheden de publieke investeringen in de vroege kinderjaren snel aan het uitbreiden; in sommige landen is al bijna sprake van universele en hoogwaardige voorzieningen. Belangrijke kwaliteitsaspecten die beleidsmakers kunnen beïnvloeden hebben vooral te maken met het soort kinderopvang, de toegepaste didactiek, en de beschikbaarheid van voldoende (gekwalificeerd) personeel. Echter, de relevantie van dergelijke aspecten voor de ontwikkeling van kinderen, en voor het tegengaan van ongewenste verschillen in ontwikkeling, hebben tot nu toe minder aandacht gekregen. Dit proefschrift biedt nieuwe bewijzen over die kwaliteitsaspecten op basis van empirische analyses van drie unieke databestanden, waarvan er twee voor het eerst beschikbaar zijn gemaakt voor onderzoek. Twee situaties worden onderzocht: Nederland, een land waar de formele kinderopvang in de afgelopen jaren snel is gegroeid, en Denemarken, een land met een lange geschiedenis van universeel hoogwaardige kinderopvang.

De eerste studie (hoofdstuk II) behandelt de effecten van kinderopvang in centra ten opzichte van een thuisomgeving; er wordt gebruik gemaakt van uitgebreide enquêtegegevens uit een grootschalig cohortonderzoek naar kinderen die in

Nederland opgroeiden aan het begin van het nieuwe millennium. Nederland heeft in de laatste twee decennia aanzienlijke investeringen gedaan om de markt voor opvang in centra te professionaliseren en het bereik te vergroten, en om kinderen beter voor te bereiden op het basisonderwijs. De studie analyseert de invloed van opvang in een centrum in plaats van thuis op een reeks ontwikkelingsresultaten van kinderen in de begintijd van de basisschool. Er wordt rekening gehouden met mogelijke niet-lineariteit in de correlatie tussen de ervaren kinderopvang en ontwikkelingsresultaten. De resultaten van schattingen naar de kleinste-kwadratenmethode en gevoeligheidsanalyses, dat wil zeggen benaderingen middels een instrumentele variabele en 'structural equation modeling', bieden geen ondersteuning aan de aanzienlijke kortetermijneffecten van centrumgebaseerde kinderopvang die worden vermeld in de literatuur.

De tweede studie (hoofdstuk III) onderzoekt of kinderen het beter doen op school als ze deelnemen aan een peuterschool met een kindgerichte in plaats van een docent-gebaseerde benadering van vroeg onderwijs, of helemaal geen onderwijsbenadering. Van voor- en vroegschoolse educatie wordt vooral verwacht dat deze kansarme kinderen helpt. Er is echter onenigheid over de vraag of het onderwijs op die leeftijd vooral gebaseerd moet zijn op de docent en het curriculum, of bij voorkeur gericht moet zijn op de ontwikkeling van het individuele kind. De studie kijkt in het bijzonder naar kindgerichte onderwijsprogramma's (overwegend 'Speelplezier') toegepast in Zuid-Limburg. De studie biedt de eerste analyse van een uitgebreide dataset, verzameld in 2008/09, die onlangs ingeschreven kinderen op basisscholen in die regio beslaat. Verschillende schattingstechnieken laten zien dat kindgerichte voor- en vroegschoolse educatie nog niet heeft geleid tot de verwachte significante effecten op resultaten aan het begin van het basisonderwijs.

De derde studie (hoofdstuk IV) gaat in op de langetermijneffecten van universele en hoogwaardige voorzieningen voor peuters en kleuters, in het bijzonder de rol van de kwaliteitsaspecten van het personeel, op de prestaties van kinderen aan het einde van de middelbare school (leeftijd 16). Een unieke combinatie van Deense administratieve bestanden maakt het mogelijk om kinderen te koppelen aan hun voorschoolse voorzieningen, en de vijf belangrijkste kwaliteitsindicatoren te genereren over het personeel van deze voorzieningen (personeel per kind; personeelsverloop; en het aandeel van professionele, mannelijke en allochtone medewerkers). Corrigerend voor achtergrondfactoren van het kind onderzoekt de studie effecten van deze indicatoren op langetermijnresultaten, en de heterogeniteit van dergelijke effecten. De verschillende analyses blijken bescheiden langetermijneffecten van een betere kwaliteit van het voorschoolse personeel op de taalontwikkeling te tonen. Vier van de vijf kwaliteitsindicatoren worden geassocieerd met significante verbeteringen in Deense taaltoetsscores - zij het dat de relatie numeriek bescheiden is. Jongens hebben meer baat van betere

voorschoolse kwaliteit dan meisjes, en allochtone kinderen profiteren van een hogere personeelscontinuïteit. Een gevoeligheidsanalyse, een aanpak die gebruik maakt van een instrumentele variabele, adresseert mogelijke onzuiverheden die kunnen voortvloeien uit onopgemerkte variabelen in de schatting naar de kleinste-kwadratenmethode; deze produceert echter geen consistentere schattingen.

De uitkomsten van deze onderzoeken geven aan dat betere kwaliteit tot positieve resultaten lijkt te leiden. De nieuwe bewijzen zijn echter niet voldoende om de hoge verwachtingen met betrekking tot voorschoolse investeringen te ondersteunen: daarvoor zijn te veel van de geschatte effecten te klein in omvang of niet-significant. Het bewezen sociale mobilisatiepotentieel van dergelijke investeringen komt evenmin overeen met de hoge verwachtingen. Een aantal mogelijke oorzaken en kritische beperkingen van de analyses wordt naar voren gebracht bij de bespreking van het bewijsmateriaal. Op basis van wat we tot nu toe weten, moeten investeringen in opvang en educatie voor de vroege kinderjaren in de eerste plaats gestoeld zijn op de wens om de arbeidsmarktparticipatie van moeders mogelijk te maken. Investeringen in specifieke kwaliteitaspecten kunnen alleen gebaseerd zijn op theoretisch of stapsgewijs bewijs. Ondanks een sterk en wijdverbreid geloof dat kwaliteit van belang is, kunnen we er nog niet zeker zijn van dat publieke middelen doeltreffend worden besteed.

Deutsche Zusammenfassung

Investitionen in Frühkindliche Betreuung und Bildung: Der Effekt von Qualität auf Ungleichheit

Diese Dissertation hat das Absicht Politiker über Investitionsmöglichkeiten frühkindlicher Betreuung und Bildung zu informieren. Sie stellt ihnen wichtige Erkenntnisse zur Verfügung in Bezug auf eine Reihe von Qualitätsmerkmalen, die es ihnen ermöglicht die Kindesentwicklung direkt positiv zu beeinflussen. In der Vergangenheit zielten öffentliche Investitionen in der frühen Kindheit in erster Linie darauf ab, die Beschäftigungsquote von Müttern zu erhöhen. Allerdings hat die Politik zunehmend erkannt, dass solche Investitionen als Instrumente einer Verbesserung sozialer Mobilität von Kindern dienen können, wenn diese Gefahr laufen, in der Schule den Anschluss zu verlieren, zum Beispiel weil sie in armen Haushalten, mit alleinerziehenden oder schlechter ausgebildeten Eltern aufwachsen, sie einer ethnischen Minderheit angehören, oder Jungen sind. Die Bedeutung einer Mobilisierung des gesamten Entwicklungspotenzials von Kindern wird durch die Tatsache hervorgehoben, dass – trotz wirtschaftlichen Fortschritts – Chancengleichheit in den meisten Länder noch nicht erreicht wurde, während zugleich die demografische Entwicklung dazu führt, dass junge Generationen in Zukunft voraussichtlich mehr belastet werden. Gleichzeitig hat frühkindliche Forschung das Argument verstärkt, dass im frühen Alter investierte öffentliche Mittel eine nachgreifende, positive Entwicklungsdynamik befördern können, die zugleich hohe Renditen für die Gesellschaft als Ganzes versprechen.

Regierungen vieler Länder sind damit beschäftigt öffentlichen Investitionen in frühkindliche Betreuung und Bildung rasch auszuweiten; einige wenige haben bereits ein nahezu universelles hochwertiges Angebot erreicht. Wichtige Qualitätsaspekte, die die Politik beeinflussen kann, sind insbesondere mit der Art der Kinderbetreuung, dem angewandten pädagogischen Ansatz, und der Verfügbarkeit adäquater Mitarbeiter verknüpft. Allerdings hat die Bedeutung dieser Aspekte für die Entwicklung des Kindes und der Kompensation unerwünscht heterogener Entwicklungsergebnisse bisher nur wenig Aufmerksamkeit erhalten. Diese Dissertation bietet neue wissenschaftliche Beweise zur Bedeutung dieser Qualitätsaspekte, basierend auf empirischen Untersuchungen dreier einzigartiger Datensätze, von denen zwei zum ersten Mal der Forschung zugänglich sind. Zwei Länder werden genauer untersucht - die Niederlande, einem Land, in dem das Angebot formaler Kinderbetreuung in den letzten Jahren rasch gewachsen ist, und Dänemark, einem Land mit einer langen Geschichte universaler und qualitativ hochwertiger Kinderbetreuung.

Die erste Studie (Kapitel II) beschäftigt sich mit den Effekten einer häuslichen verglichen mit einer institutionellen Kinderbetreuung, unter Verwendung eines reichhaltigen Datensatzes einer Kohorte von Kindern, die in den Niederlanden zu Beginn des neuen Jahrtausends aufgewachsen sind. Die Niederlande haben in den letzten zwei Jahrzehnten erhebliche Investitionen getätigt, um das Betreuungsangebot auszubauen und zu professionalisieren, und um die gegebene Schulreife zu verbessern. Die Studie analysiert die Auswirkungen häuslicher bzw. institutioneller Kinderbetreuung auf eine Reihe von Entwicklungsergebnissen zu Beginn der Grundschulzeit. Die Studie berücksichtigt etwaige Nichtlinearität in der Korrelation zwischen erfahrener Kinderbetreuung und der Kindesentwicklung. Ergebnisse von Schätzungen nach der Kleinstquadrat-Methode und von Sensitivitätsanalysen, in diesem Falle die Anwendung von instrumentellen Variablen und von Strukturgleichungsmodellen, unterstützen nicht die Aussage erheblicher kurzfristiger Effekte die einer institutionellen Kinderbetreuung in der Literatur zugesprochen werden.

Die zweite Studie (Kapitel III) untersucht, ob Kinder besser in der Schule sind, wenn sie eine Vorschule besucht haben, die einen Kind-zentrierten Ansatz statt eines Lehrer-basierten bzw. keines dergleichen zur Früherziehung angewandt haben. Vor frühkindlicher Bildung wird erwartet, dass sie insbesondere benachteiligten Kindern helfen kann. Allerdings herrscht Uneinigkeit darüber, ob Bildung in diesem Alter eher Lehrer- und Lehrplan-basiert sein soll, oder vielmehr auf die individuellen Entwicklungsbedürfnisse eines einzelnen Kindes ausgerichtet sein soll. Die Studie befasst sich mit Kind-zentrierten Bildungsprogrammen (vornehmlich der Methode ‚Speelplezier‘), die im Süden von Limburg angewandt werden, einer niederländischen Provinz. Die Studie stellt die erste Analyse eines umfangreichen Datensatzes dar, der im Jahr 2008/09 erstellt wurde und Kinder erfasst, die kurz zuvor an Grundschulen in dieser Region angemeldet wurden. Verschiedene Schätzmethoden zeigen, dass Kind-zentrierte frühkindliche Bildung noch nicht die erwarteten, erheblichen Auswirkungen auf die Kindesentwicklung in der frühen Grundschulzeit erbracht hat.

Die dritte Studie (Kapitel IV) untersucht die langfristigen Vorteile von universellen, hochwertigen Vorschulen, insbesondere der Rolle von Qualitätsaspekten des Vorschulpersonals auf Entwicklungsergebnisse am Ende der Mittelschule (Alter 16). Das einzigartige Zusammenführen von dänischen Verwaltungsregistern erlaubt es Datensätze von Kindern mit denen von ihnen besuchten Vorschulen zu verknüpfen, und fünf Hauptindikatoren zu erstellen, die die Qualität des Personal von Vorschulen erfasst (Personal pro Kind; Fluktuation des Personals; der Anteil professionellen und männlichen Personals; sowie der Personalanteil von ethnischen Minderheiten). Unter Einbezug individueller Hintergrundfaktoren der Kinder, untersucht die Studie Auswirkungen dieser Indikatoren auf langfristige Entwicklungsergebnisse und auf die Heterogenität

solcher. Die verschiedenen Analysen belegen geringe, langfristige Auswirkungen einer besseren Qualität des Vorschulpersonals auf sprachliche Entwicklung. Während die betrachteten Zusammenhänge numerisch eher bescheiden sind – vier von fünf Qualitätsindikatoren zeigen signifikante Entwicklungsverbesserungen auf. Jungen profitieren mehr von besserer Vorschulqualität als Mädchen; Kinder aus ethnischen Minderheiten profitieren aus von einer höheren Kontinuität des Personals. Eine Sensitivitätsanalyse, die Anwendung instrumentaler Variablen, korrigiert mögliche Verzerrungen, die sich aus nicht beobachtete Variablen der Schätzungen nach der Kleinstquadrat-Methode ergeben; diese ermöglicht jedoch keine widerspruchsfreieren Schätzungen.

Die Ergebnisse der Studien deuten darauf hin, dass eine gute Qualität zu besseren Entwicklungsergebnissen führen kann. Allerdings verleihen die erbrachten Indizien nicht genügend Unterstützung, um die hohen Hoffnungen von Investitionen in bessere frühkindliche Betreuung und Bildung zu rechtfertigen, weil zu viele Effektschätzungen zu schwach oder nicht signifikant sind. Die Indizien erlauben es ebenfalls nicht das Potenzial einer sozialen Mobilisierung durch solche Investitionen ausreichend zu bestätigen. Eine Reihe möglicher Gründe und kritischer Einschränkungen der Analysen ergänzt die Diskussion der vorgebrachten wissenschaftlichen Beweise. Als Ergebnis dessen, was wir bisher wissen, dienen Investitionen in frühkindliche Betreuung und Bildung in erster Linie zur erleichterten Teilnahme von Müttern am Arbeitsmarkt. Investitionen in bestimmte Qualitätsaspekte können sich nur auf theoretische oder partielle Nachweise berufen. Trotz einer starken, weit verbreiteten Überzeugung, dass Qualität zählt, können wir noch nicht sicher sein, dass öffentliche Ressourcen effektiv genutzt werden.

Naložbe v varstvo in izobraževanje v zgodnjem otroštvu: Vpliv kakovosti na neenakopravnost

Ta doktorska disertacija se osredotoča na informiranje snovalcev politik o načinih investiranja v skrb in izobraževanje v zgodnjem otroštvu. Doktorska naloga prispeva pomembne dokaze o številnih značilnostih standardov kakovosti, ki jih lahko snovalci politik neposredno vključijo v spodbujanje razvoja otroka. V preteklosti so bile javne investicije v zgodnje otroštvo usmerjene predvsem v zaposlovanje mater. Vendar so te investicije snovalci politik vse bolj prepoznavali kot instrumente za stimuliranje socialne mobilnosti otrok, predvsem tistih, ki so v nevarnosti, da v šoli ne napredujejo, ker odraščajo v revnih gospodinjstvih, z enim staršem ali nižje izobraženimi starši, odraščajo v družinah, ki pripadajo narodnostni manjšini ali so dečki. Pomen mobilizacije celotnega potenciala otrok potrjuje dejstvo, da navkljub gospodarskemu napredku enakost možnosti v mnogih državah ni zagotovljena, medtem ko demografski trendi kažejo na vse večje breme, ki ga bodo morale mlade generacije nositi v prihodnosti. Obenem je raziskava zgodnjega otroštva utemeljila princip investiranja javnih sredstev v zgodnje otroštvo za pospeševanje pozitivne dinamike razvoja v nadaljnjem odraščanju, kar naj bi bilo koristno za družbo kot celoto.

Vlade v številnih državah pospešeno povečujejo javna sredstva za skrb in izobraževanje v zgodnjem otroštvu; nekatere izmed njih že dosegajo skoraj univerzalne visokokakovostne standarde. Pomembni vidiki kakovosti, na katere lahko snovalci politik vplivajo, se še posebej nanašajo na tip skrbi za otroke, uporabljen pristop izobraževanja in razpoložljivost ustrezno usposobljenih kadrov. Vseeno pa sta bila pomen tovrstnih vidikov za razvoj otrok in kompenzacija neželene heterogenosti pri uspešnosti otrok doslej deležna manj pozornosti. Ta disertacija ponuja nove dokaze o teh vidikih kakovosti na podlagi empiričnih analiz treh edinstvenih podatkovnih sklopov, od katerih sta bila dva prvič na voljo za znanstvene raziskave. Raziskovali smo dva okvirja – Nizozemska, kjer so se pogoji formalne skrbi za otroke v zadnjih letih hitro razširili ter Danska, ki je država z dolgo tradicijo univerzalnih visokokakovostnih standardov skrbi za otroke.

Prva študija (II. poglavje) se nanaša na učinke skrbi za otroke v centrih v primerjavi s tisto doma; obsežna študija s številnimi podatki je vključevala otroke, ki so odraščali na Nizozemskem na začetku tretjega tisočletja. Nizozemska je v zadnjih dveh dekadah namenila precej sredstev, da bi pogoji v izobraževalnih centrih postali bolj strokovni in obsežni ter da bi se izboljšala pripravljenost otrok na šolo. Študija analizira vpliv institucionalne skrbi za otroke na številne rezultate

otrok na začetku osnovne šole. Študija pojasnjuje potencialno nelinearnost v povezavi med izkustvom skrbi za otroke in rezultati otrok. Rezultati metode najmanjših kvadratov in analiz občutljivosti, t.j. modelni pristopi z instrumentalnimi spremenljivkami in strukturnimi enačbami, ne potrjujejo bistvenih kratkoročnih vplivov institucionalne skrbi za otroke, ki jih navaja literatura.

Druga študija (III. poglavje) se ukvarja z vprašanjem, ali so otroci uspešnejši v šoli, če obiskujejo predšolske ustanove, ki namesto učiteljskega izobraževalnega pristopa (ali celo brez tega) temeljijo na izobrazbi usmerjeni k posameznemu otroku. Izobraževanje v zgodnjem otroštvu naj bi še posebej pomagalo prikrajšanim otrokom. Vseeno pa stroka ni edina, ali naj bo izobraževanje v tem obdobju bolj osnovano na učitelju in učnem načrtu, ali pa bi se moralo osredotočati na razvoj posameznega otroka. Študija še posebej obravnava institucionalne izobraževalne programe za otroke (še posebej programa »Speelplezier«), ki ga izvajajo v nizozemski provinci Limburg. Predstavljena je prva analiza obsežnega podatkovnega sklopa iz 2008/09, ki vključuje v zadnjem času vpisane otroke v osnovne šole na tem področju. Različne tehnike ocenjevanja kažejo, da zgodnje izobraževanje, osredotočeno na otroka, še ni privedlo do pričakovanih bistvenih učinkov na uspešnost otroka v nižjih razredih osnovne šole.

Tretja študija (IV. poglavje) preučuje dolgoročne koristi univerzalno visokokakovostnih standardov predšolske vzgoje, še posebej vlogo kakovosti predšolskih delavcev na uspeh otrok na koncu osnovne šole (starost 16 let). Edinstvena skladnost danskih upravnih registrov omogoča, da se poveže otroke z njihovim predšolskim izobraževanjem in tako ustvari pet glavnih kazalnikov, ki opisujejo kakovost izobraževalnih delavcev v teh predšolskih programih (en delavec na otroka; fluktuacija kadrov ter delež strokovnih in moških kadrov in kadrov iz narodnostnih manjšin). Študija za nadzor dejavnikov okolja otrok razišče vplive teh kazalnikov na dolgoročen uspeh in heterogenost teh učinkov. Različne analize kažejo na skromne dolgoročne učinke kakovostnejših predšolskih delavcev na jezikovni razvoj otrok. Štirje od petih kazalnikov kakovosti so povezani z bistvenim izboljšanjem rezultatov testov danskega jezika, ki so jih opravili otroci na Danskem – čeprav je razmerje številčno skromno. Dečki več pridobijo od kakovostnejših predšolskih programov kot deklice, medtem ko otrokom iz narodnostnih manjšin koristi večja stabilnost kadrov. Analiza občutljivosti, oziroma pristop z instrumentalno spremenljivko, kaže na morebitno pristranskost, ki lahko izhaja iz neupoštevanih spremenljivk v metodi najmanjših kvadratov; vseeno pa ne prinaša bolj konsistentnih ocen.

Rezultati teh študij kažejo, da lahko večja kakovost pripelje do pozitivnih rezultatov. Vseeno pa novi dokazi niso dovolj, da bi potrdili in upravičili velika upanja investiranja v skrb za otroke, saj je preveč ocen učinkov prenizkih oziroma

nebistvenih, poleg tega pa se dokazi potencialne socialne mobilizacije tovrstnih investicij ne ujemajo z visokimi pričakovanji. V razpravi dokazov so predstavljeni številni možni razlogi in kritične omejitve analiz. Iz tega lahko sklepamo, da so sredstva za skrb in izobraževanje v zgodnjem otroštvu v prvi vrsti odvisna od motivacije za pospešitev sodelovanja mater na trgu dela. Investicije v specifične vidike kakovosti je možno utemeljiti na teoretičnih oziroma posameznih dokazih. Kljub močnemu in širokem prepričanju, da je kakovost pomembna, pa ne moremo z gotovostjo govoriti o učinkoviti porabi javnih sredstev.

Curriculum Vitae

Mr Robert Bauchmüller pursued higher education at Maastricht University, the Netherlands. Studying international economics, he followed the track Human Resources, Labour Markets and Society. In 2005, he was awarded the Dutch equivalent degree to an MSc in Economics. During his graduate studies, he studied for an Erasmus semester at the Warsaw School of Economics, and did an AIESEC Development Internship at the NGO Sneh Prayas in Ahmedabad, India. And he was active in a number of student organisations, such as the AEGEE Network and the Model European Union Foundation, which promote exchange and cooperation among young European citizens; he also served the latter as president.

In 2002-2010, Mr Bauchmüller worked in several positions at Maastricht University – first at the School of Business and Economics and later at the Graduate School of Governance. He worked, for example, for the Master Programmes in Social Protection Financing & Policy Design, which are organised in cooperation with the ILO, and was involved in a number of consultancy and training projects. He was also Programme Manager of the first UNICEF Learning Programme on Public Policy, Advocacy and Partnerships for Children’s Rights.

In 2006-2010, Mr Bauchmüller was a Research Fellow in the PhD Programme on Public Policy & Policy Analysis, for which he fulfilled an extensive course requirement. He received additional training at a Summer Workshop of the Social Change Harvard-Manchester Initiative, the Essex Summer School in Social Science and Data Analysis, and the Barcelona Micro-econometrics Summer School. He made research visits to the Institute for Social & Economic Research at the University of Essex, UK, the Department of Economics, Business & Statistics at the University of Milan, Italy, and the Danish Institute for Local and Regional Government Research in Copenhagen, Denmark. He attended and presented his research at several international conferences and seminars.

Since 2011, Mr Bauchmüller has been employed at the Center of Excellence in Finance (CEF) in Ljubljana, Slovenia, which responds to capacity development needs in public financial management and central banking of countries across South East Europe. Upon joining the CEF, he has been in charge of the CEF’s multi-beneficiary programme for Building Capacities in Policy Design and Implementation. Since recently, he leads the team of the Secretariat that the CEF provides to the Public Expenditure Management Peer Assisted Learning (PEMPAL) network, a multilateral effort to develop capacity and share reform experiences among countries in Central Asia and Central and East Europe.

Maastricht Graduate School of Governance Dissertation Series

Rehm, Martin

Unified yet separated:

Empirical Study on the Impact of Hierarchical Positions within Communities of Learning

MGSOG Dissertation Series, # 32 (2013)

Mbuvi, Dorcas

Utility reforms and performance of the Urban Water Sector in Africa

MGSOG Dissertation Series, # 31 (2012)

Salanauskaite, Lina

Essays on the Distributional Impact of Public Policies, Ex-Ante and Ex-Post Evaluation

MGSOG Dissertation Series, # 30 (2012)

Schuering, Esther

To condition or not - is that the question? An analysis of the effectiveness of ex-ante and ex-post conditionality in social cash transfer programs

MGSOG Dissertation Series, # 29 (2012)

Abah, Joe

Strong Organisations in Weak States

MGSOG Dissertation Series, # 28 (2012)

Nimeh, Zina Samih

Social citizenship rights: Inequality and exclusion

MGSOG Dissertation Series, # 27 (2012)

Eisenhamerova, Lenka

Legitimacy of 'humanitarian military intervention'

MGSOG Dissertation Series, # 26 (2011)

Tomini, Sonila

Informal Payments for Health Care Services in Albania

MGSOG Dissertation Series, # 25 (2011)

Li, Jinjing

Dynamic Microsimulation for Public Policy Evaluation

MGSOG Dissertation Series, # 24 (2011)

Atamanov, Aziz

Rural Nonfarm Employment and International Migration as Alternatives to Agricultural Employment: The Case of Kyrgyzstan

MGSOG Dissertation Series, # 23 (2011)

Vandeninden, Frieda

Poverty Alleviation: Aid and Social Pensions

MGSOG Dissertation Series, # 22 (2011)

Tirivayi, Nyasha

The Welfare Effects of Integrating AIDS Treatment with Food Transfers: Evidence from Zambia

MGSOG Dissertation Series, # 21 (2011)

Sowa, Agnieszka

Who's left behind?

Social dimensions of health transition and utilization of medical care in Poland

MGSOG Dissertation Series, # 20 (2011)

Sfakianakis, Emmanouil

The Role of Private Actors in the Provision of Public Goods with Applications to Infrastructure and Financial Stability

MGSOG Dissertation Series, # 19 (2011)

Lo, Siu Hing

White Collars Green Sleeves: An Interorganizational Comparison of Determinants of Energy-Related Behaviors among Office Workers

MGSOG Dissertation Series, # 18 (2011)

Wu, Treena

Constraints to Human Capital Investment in Developing Countries:

Using the Asian Financial Crisis in Indonesia as a Natural Experiment

MGSOG Dissertation Series, # 17 (2011)

Espinoza Pena, Henry

Impact Evaluation of a Job-Training Programme for Disadvantaged Youths: The Case of Projoven

MGSOG Dissertation Series, # 16 (2011)

Tomini, Florian

Between family and friends: Understanding the interdependency of private transfers

MGSOG Dissertation Series, # 15 (2010)

Polalowski, Michał

*The Institutional Transformation of Social Policy in East Central Europe:
Poland and Hungary in comparative and historical perspective*

MGSOG Dissertation Series, # 14 (2010)

Ahmed, Maha

*Defining, Measuring and Addressing Vulnerability:
The Case of Post-Conflict Environments*

MGSOG Dissertation Series, # 13 (2010)

Beckers, Pascal

*Local space and economic success:
The role of spatial segregation of migrants in the Netherlands*

MGSOG Dissertation Series, # 12 (2010)

Cebotari, Victor

*Conflicting Demands in Ethnically Diverse Societies Ethnopolitical Contention and
Identity Values in Europe*

MGSOG Dissertation Series, # 11 (2010)

Gyllensporre, Dennis

*Competing and Complementarity Perspectives on the EU as a Crisis Actor Management
Actor: An Examination of the Common Security and Defence Policy through the Lenses of
Idealism and Realism*

MGSOG Dissertation Series, # 10 (2010)

Vall Castello, Judit

*Business Cycle and Policy Effects on Labour Market Transitions of Older and Disabled
Workers in Spain*

MGSOG Dissertation Series, # 9 (2010)

Roelen, Keetie

*False Positives or Hidden Dimensions:
The Definition and Measurement of Child Poverty*

MGSOG Dissertation Series, # 8 (2010)

Sologon, Denisa Maria

Earnings Dynamics in Europe

MGSOG Dissertation Series, # 7 (2010)

Siegel, Melissa

Money and Mobility: Migration and Remittances

MGSOG Dissertation Series, # 6 (2010)

Hagen-Zanker, Jessica Sabine

Modest expectations:

Causes and effects of migration on migrant households in source countries

MGSOG Dissertation Series, # 5 (2010)

Muñiz Castillo, Mirtha R.

Human Development and Autonomy in Project Aid:

Experiences from four bilateral projects in Nicaragua and El Salvador

MGSOG Dissertation Series, # 4 (2009)

Arndt, Christiane

Governance Indicators

MGSOG Dissertation Series, # 3 (2009)

Augsburg, Britta

Microfinance - Greater Good or Lesser Evil?

MGSOG Dissertation Series, # 2 (2009)

Notten, Geranda

Measuring and Managing Poverty Risks

MGSOG Dissertation Series, # 1 (2008)

